Wolfgang Löffelhardt *Editor*

Endosymbiosis

Springer

Endosymbiosis

Wolfgang Löffelhardt
Editor

Endosymbiosis

Springer

Editor
Wolfgang Löffelhardt
Department of Biochemistry
University of Vienna
Vienna
Austria

ISBN 978-3-7091-1302-8 ISBN 978-3-7091-1303-5 (eBook)
DOI 10.1007/978-3-7091-1303-5
Springer Wien Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013953936

© Springer-Verlag Wien 2014
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Endosymbiosis is a key process in the evolution of the eukaryotic cells and thus a central theme in biology. The approach chosen in this book puts emphasis and focus on the plant kingdom. The driving force for plastid endosymbiosis was the gain of autotrophy with photosynthesis as the base for higher forms of life on our planet. A eukaryotic host cell engulfed a cyanobacterium or a eukaryotic alga resulting after long-lasting and highly complex adaptations in phototrophic organisms harboring primary, secondary, or tertiary plastids, respectively. In the case of mitochondria, no candidate for an anaerobic eukaryotic host cell that would engulf an α-proteobacterium has yet been found and hypotheses involving merging of two prokaryotes, an archaebacterium and an α-proteobacterium, received attention in the past years. Franz Lang gives a balanced view on these possibilities.

My own chapter tries to convince you about the single primary endosymbiotic event and the monophyly of the kingdom "Plantae." Good evidence for that comes from the recently sequenced genome of *Cyanophora* (the peptidoglycan-surrounded muroplasts being the biochemical proof for cyanobacterial ancestry) as put forward by Bhattacharya and colleagues. The acquisition of metabolite transporters (Facchinelli and Weber) and the concerted merging of host and endosymbiont reserve carbohydrate biosynthesis and degradation (Steven Ball) are crucial for the successful onset of an endosymbiosis. The evolution of plastid protein import (Sommer and Schleiff) is juxtaposed to the second chapter on mitochondria (Hewitt and colleagues) dealing with the same issue. It is amazing how similar apparatus could be created from a quite different set of subunits.

Also, intermediate stages of organelle evolution are better documented at present for plastids than for mitochondria: Yoon and colleagues describe the "plastid in the making" of *Paulinella* and Adler and colleagues a novel nitrogen-fixing organelle-to-be of *Rhopalodia*. Another asset of plastids is the complex plastids that have no parallels in mitochondria, where no secondary and tertiary endosymbioses are known. Three chapters deal with several aspects of secondary endosymbioses: Tanifuji and Archibald depict the nucleomorph, the vestigial nucleus of the primary host cell (and as such the biochemical proof for the correctness of the model) in cryptomonads and chlorarachnids; Grosche and colleagues illustrate the problems

encountered by plastid-targeted proteins in complex plastids; Linares and colleagues present a novel organism, *Chromera*, the first phototrophic relative to apicomplexans. Tertiary endosymbioses are dealt with by Gagat and colleagues in a comprehensive chapter on dinoflagellates with their various types of plastids, genuine, exchanged, or stolen. These superimposed secondary and tertiary endosymbioses are mainly responsible for algal biodiversity and speciation. Finally, Wägele and Martin answer the long-prevailing question of a tertiary endosymbiosis between animals (sea slugs) and algae to the negative: the longevity of plastids in *Elysia* is striking—but they are just kleptoplastids.

Wien, July 2013 Wolfgang Löffelhardt

Contents

Part I The Heterotrophic Eukaryotes

Mitochondria and the Origin of Eukaryotes 3
B. Franz Lang

Modifications and Innovations in the Evolution of Mitochondrial Protein Import Pathways 19
Victoria Hewitt, Trevor Lithgow, and Ross F. Waller

Part II Autotrophy as the Driving Force for Endosymbiosis: Primary Endosymbiosis

The Single Primary Endosymbiotic Event 39
Wolfgang Löffelhardt

Insertion of Metabolite Transporters into the Endosymbiont Membrane(s) as a Prerequisite for Primary Endosymbiosis 53
Fabio Facchinelli and Andreas P.M. Weber

Evolution of the Protein Translocons of the Chloroplast Envelope 81
Maik S. Sommer and Enrico Schleiff

Evolution of Storage Polysaccharide Metabolism in Archaeplastida Opens an Unexpected Window on the Molecular Mechanisms That Drove Plastid Endosymbiosis 111
S.G. Ball

Analysis of the Genome of *Cyanophora paradoxa*: An Algal Model for Understanding Primary Endosymbiosis 135
Debashish Bhattacharya, Dana C. Price, Cheong Xin Chan, Jeferson Gross, Jürgen M. Steiner, and Wolfgang Löffelhardt

Part III Recent "Primary" Endosymbioses

Photosynthetic *Paulinella*: Recapitulation of Primary Plastid Establishment 151
Hwan Su Yoon, Eun Chan Yang, Huan Qiu, and Debashish Bhattacharya

***Rhopalodia gibba*: The First Steps in the Birth of a Novel Organelle?** 167
Susann Adler, Eike M. Trapp, Christine Dede, Uwe G. Maier, and Stefan Zauner

Part IV Autotrophy as the Driving Force for Endosymbiosis: Secondary and Tertiary Endosymbioses

***Chromera* et al.: Novel Photosynthetic Alveolates and Apicomplexan Relatives** 183
Marjorie Linares, Dee Carter, and Sven B. Gould

Nucleomorph Comparative Genomics 197
Goro Tanifuji and John M. Archibald

Protein Import into Complex Plastids: Current Findings and Perspectives 215
Christopher Grosche, Franziska Hempel, Kathrin Bolte, Lars Abram, Uwe G. Maier, and Stefan Zauner

Tertiary Plastid Endosymbioses in Dinoflagellates 233
Przemysław Gagat, Andrzej Bodył, Paweł Mackiewicz, and John W. Stiller

Endosymbioses in Sacoglossan Seaslugs: Plastid-Bearing Animals that Keep Photosynthetic Organelles Without Borrowing Genes 291
Heike Wägele and William F. Martin

Index 325

List of Contributors

Lars Abram Laboratorium für Zellbiologie, Philipps-Universität Marburg, Marburg, Germany

Institut für Zell- und Entwicklungsbiologie der Pflanzen, Universität Bielefeld, Bielefeld, Germany

Susann Adler Laboratorium für Zellbiologie, Philipps-Universität Marburg, Marburg, Germany

John M. Archibald Canadian Institute for Advanced Research, Integrated Microbial Biodiversity Program, Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, NS, Canada

S.G. Ball Unité de Glycobiologie Structurale et Fonctionnelle, Université des Sciences et Technologies de Lille, Lille, France

Debashish Bhattacharya Rutgers University, New Brunswick, NJ, USA

Department of Ecology, Evolution and Natural Resources, Rutgers University, New Brunswick, NJ, USA

Andrzej Bodył Laboratory of Evolutionary Protistology, Department of Evolutionary Biology and Ecology, University of Wrocław, Wrocław, Poland

Kathrin Bolte Laboratorium für Zellbiologie, Philipps-Universität Marburg, Marburg, Germany

Dee Carter School of Molecular Bioscience, University of Sydney, NSW, Australia

Cheong Xin Chan Department of Ecology, Evolution and Natural Resources, Institute of Marine and Coastal Science, Rutgers University, New Brunswick, NJ, USA

The University of Queensland, Institute for Molecular Bioscience and ARC Centre of Excellence in Bioinformatics, Brisbane, QLD, Australia

Christine Dede Laboratorium für Zellbiologie, Philipps-Universität Marburg, Marburg, Germany

Fabio Facchinelli Institut für Biochemie der Pflanzen, Heinrich-Heine Universität Düsseldorf, Düsseldorf, Germany

Przemysław Gagat Department of Genomics, Faculty of Biotechnology, University of Wrocław, Wrocław, Poland

Sven B. Gould Institute for Molecular Evolution, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany

Christopher Grosche Laboratorium für Zellbiologie, Philipps-Universität Marburg, Marburg, Germany

Jeferson Gross Department of Ecology, Evolution and Natural Resources, Institute of Marine and Coastal Science, Rutgers University, New Brunswick, NJ, USA

Franziska Hempel Laboratorium für Zellbiologie, Philipps-Universität Marburg, Marburg, Germany

LOEWE-Zentrum für Synthetische Mikrobiologie (SynMikro), Marburg, Germany

Victoria Hewitt Department of Biochemistry and Molecular Biology, Monash University, Clayton Campus, Melbourne, Australia

B. Franz Lang Robert Cedergren Centre of Bioinformatics and Genomics, Department of Biochemistry, Université de Montréal, Montréal, QC, Canada

Marjorie Linares School of Molecular Bioscience, University of Sydney, NSW, Australia

Trevor Lithgow Department of Biochemistry and Molecular Biology, Monash University, Clayton Campus, Melbourne, Australia

Wolfgang Löffelhardt Max F. Perutz Laboratories, Department of Biochemistry and Cell Biology, University of Vienna, Vienna, Austria

Paweł Mackiewicz Department of Genomics, Faculty of Biotechnology, University of Wrocław, Wrocław, Poland

Uwe G. Maier Laboratorium für Zellbiologie, Philipps-Universität Marburg, Marburg, Germany

LOEWE-Zentrum für Synthetische Mikrobiologie (SynMikro), Marburg, Germany

William F. Martin Institut für Molekulare Evolution, Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany

Dana C. Price Department of Ecology, Evolution and Natural Resources, Institute of Marine and Coastal Science, Rutgers University, New Brunswick, NJ, USA

Huan Qiu Bigelow Laboratory for Ocean Sciences, West Boothbay Harbor, ME, USA

Enrico Schleiff Cluster of Excellence Macromolecular Complexes, Centre of Membrane Proteomics and Institute for Molecular Biosciences, Department of Biosciences, Molecular Cell Biology of Plants, Goethe University, Frankfurt, Germany

Maik S. Sommer Cluster of Excellence Macromolecular Complexes, Centre of Membrane Proteomics and Institute for Molecular Biosciences, Department of Biosciences, Molecular Cell Biology of Plants, Goethe University, Frankfurt, Germany

Jürgen M. Steiner Department of Biology/Plant Physiology, Martin-Luther-University Halle-Wittenberg, Halle (Saale), Germany

John W. Stiller Department of Biology, East Carolina University, Howell Science, Greenville, NC, USA

Goro Tanifuji Canadian Institute for Advanced Research, Integrated Microbial Biodiversity Program, Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, NS, Canada

Eike M. Trapp Laboratorium für Zellbiologie, Philipps-Universität Marburg, Marburg, Germany

Heike Wägele Zoologisches Forschungsmuseum Alexander Koenig, Bonn, Germany

Ross F. Waller School of Botany, University of Melbourne, Melbourne, Australia

Andreas P.M. Weber Institut für Biochemie der Pflanzen, Heinrich-Heine Universität Düsseldorf, Düsseldorf, Germany

Eun Chan Yang Bigelow Laboratory for Ocean Sciences, West Boothbay Harbor, ME, USA

Hwan Su Yoon Department of Biological Sciences, Sungkyunkwan University, Suwon, South Korea

Stefan Zauner Laboratorium für Zellbiologie, Philipps-Universität Marburg, Marburg, Germany

Part I
The Heterotrophic Eukaryotes

Mitochondria and the Origin of Eukaryotes

B. Franz Lang

Abstract The hypothesis that mitochondria evolved via endosymbiosis of an α-proteobacterium with a primitive eukaryotic host cell has long become part of modern textbooks. Yet, many fundamental questions remain unanswered, such as the nature of the host, the evolution of distinctive eukaryotic features, and the events leading to the transition from a symbiont to an organelle. This is because of the limited phylogenetic evidence for ancient evolutionary events, leaving these questions open to speculation. According to the serial endosymbiont hypothesis, eukaryotes evolved gradually from an ancient lineage without mitochondria (termed "archezoa"), and at some later point acquired mitochondria by endosymbiosis with an α-proteobacterium. More recent theories posit that eukaryotes (E) originated by metabolic symbiosis (syntrophy) of an archaeal (a) with a bacterial cell (b). A merger of two simple prokaryotic organisms (E = a + b) would give rise to the radically different eukaryotic cell as well as to mitochondria. The underlying tenet of these theories is that evolution proceeds from simple (primitive) to complex ("higher" in an Aristotelian sense) organisms and that the symbiotic event itself somehow leads to the rapid emergence of highly complex subcellular structures such as nucleus and cytoskeleton.

This review will discuss theories on the origin of mitochondria and the eukaryotic cell and confront them with results from rigorous phylogenetic inferences. One of the strongest arguments against symbiogenesis scenarios is that complex subcellular structures and thousands of genes would have had to emerge within a short evolutionary interval and without leaving any record of intermediate forms.

B.F. Lang (✉)
Robert Cedergren Centre of Bioinformatics and Genomics, Department of Biochemistry, Université de Montréal, Montréal, QC, Canada
e-mail: Franz.Lang@Umontreal.ca

W. Löffelhardt (ed.), *Endosymbiosis*, DOI 10.1007/978-3-7091-1303-5_1,

© Springer-Verlag Wien 2014

Contents

Introduction 4
The Evolution of Eukaryotes Is One of Symbiotic Relationships 4
A Single Origin of Mitochondria? 6
The Archezoan (SET) Hypothesis on the Origin of Eukaryotes and Mitochondria 7
A Eukaryotic Origin by Symbiogenesis? 8
Implications of the Hydrogen Hypothesis 9
Conclusions 13
References 15

Introduction

The origin of the eukaryotic cell continues to be among the most controversial questions in evolutionary biology. This is because key events leading to the divergence of Eukarya, Bacteria, and Archaea date back a few billion years ago, and therefore traces left in genomes of extant species are difficult to interpret. This issue is amplified by repeated periods of massive species extinction, which in phylogenetic inference leaves us with uneven species sampling and in some cases lack of an appropriate outgroup. As a consequence, theories on the origin of eukaryotes have so far been highly speculative, with some authors even explicitly disregarding phylogenetic results. Yet, given the quickly growing number of known genome sequences and much improved phylogenetic methodology, testing for potential incoherencies in current theories on mitochondrial and eukaryotic origins is in order and has become feasible.

In this review, we will revisit the question whether mitochondria emerged indeed only once, and how traditional theories on the origin of eukaryotes and mitochondria hold up against more recent "symbiogenesis" theories. In addition, we will discuss recent phylogenetic analyses on the rooting of the eukaryotic tree, the origin of archaea-related eukaryotic genes, and the impact of these analyses on our understanding of early eukaryotic evolution.

The Evolution of Eukaryotes Is One of Symbiotic Relationships

Among the three domains of cellular organisms, Bacteria, Archaea, and Eukarya, the latter stands out by a phagocytotic mode of nutrition (the majority of protists are microbe-eating predators) and numerous symbiotic relationships. For instance, animals require an elaborate population of diverse, intestinal microbes; most plants rely on mycorrhizal fungi and soil bacteria to thrive; and many unicellular eukaryotes, in particular amoebae, carry (often proteo-) bacterial intracellular symbionts. The propensity for complex and persistent symbiotic relationships is

an ancient eukaryotic trait, as witnessed by their organelles of α-proteobacterial and cyanobacterial origin (mitochondria and plastids, respectively).

The interplay between eukaryotes and bacterial endosymbionts can be highly complex and intimate. For instance, the ciliate *Paramecium* captures *Holospora obtusa* (a distant relative of rickettsial pathogens) as bacterial food; yet, in contrast to other bacteria, *Holospora* remains undigested in the food vacuole and migrates via the cytoplasm into the host's macronucleus (e.g., Fujishima et al. 1997; Fujishima and Fujita 1985; Görtz et al. 1990; Wiemann and Görtz 1991). Note that this migration implies crossing of two membrane barriers, which involves the participation of a multitude of specialized bacterial and ciliate proteins and functions. The bacteria multiply and mature within the nucleus and eventually depart to start a new infection cycle—surprisingly without apparent damage to the host cell. Due to their specialized life style, bacterial endosymbionts undergo rapid reductive genome evolution and an increasing reliance on host functions (Wixon 2001). This also applies to *Holospora* with a genome size of less than 2 Mbp, and lack of ATP production via oxidative phosphorylation (Lang et al. 2005). Examples like this demonstrate that eukaryotes are receptive or at least tolerant to accommodating intracellular guests, dedicating various complex subcellular structures to predation and to hosting of endosymbionts.

Eukaryotes not only team up with partners from the other domains of life but also among themselves. For example, certain algae plus fungi combine into lichens, a most resilient form of life, and non-photosynthetic protists unite with algae, which has led in several instances to secondary endosymbiosis, and when repeated, tertiary and higher order endosymbioses. In an extreme case, five distinct genomes (i.e., two nuclear, two mitochondrial, and one plastid) have been identified in cells of an organism that is the product of a relatively recent union of a dinoflagellate with a diatom (Imanian et al. 2010). The benefit for the symbiotic host seems evident in the case of mitochondria and plastids, which equip the host with oxidative phosphorylation and photosynthesis, among other features. In other cases, a benefit is uncertain (e.g., *Holospora* symbionts) or simply absent (pathogenic Rickettsiales symbionts; phylogenetic relatives of *Holospora* and mitochondria). Ancient symbiotic events are often viewed as driven by selective advantage for one or both partners, but it is equally possible that they have started out by pure accident, or even as consequence of a microbial disease.

The above examples make clear that the evolution of eukaryotes is one of symbiotic relationships, implying that the genetic material of extant eukaryotic lineages may come from a wide variety of sources, which makes prediction of mitochondrial and eukaryotic origins a most challenging task. For this, thorough phylogenetic analysis is required, even more so because of known repeated associations with similar partners (e.g., mitochondria, followed or preceded by intracellular endosymbioses with other α-proteobacteria). Another important point is that although symbiotic events may be favored through the acquisition of beneficial features by one or the other partner, there may be no tangible advantage other than having survived an act of "biological warfare."

With the strong propensity of eukaryotes to engage over again in new symbiotic relationships, the question arises if mitochondria have indeed emerged only once as widely assumed and reconfirmed. Do we have clear-cut phylogenetic evidence that makes multiple bacterial acquisitions improbable?

A Single Origin of Mitochondria?

Mitochondria have become a central issue in discussing eukaryotic origins, as apparently all extant eukaryotic lineages either have or once had mitochondria. The emphasis is on "apparently," because this view is based on the assumption that mitochondria, hydrogenosomes, and mitosomes originate from the very same symbiotic event. But with an estimated origin of extant eukaryotes close to a billion or more years ago (Brinkmann and Philippe 2007) and only a few proteins left in hydrogenosomes and mitosomes that are clearly of α-proteobacterial origin, what is the accuracy of this statement? How well may we distinguish between one or more ancient symbiotic events with members from within a given bacterial group? The somewhat sobering answer is that single- and multi-gene phylogenies alike have difficulties to even distinguish between a sister group relationship of mitochondria and *Rickettsia*-related bacteria, versus an origin from within the *Rickettsia* lineage or within free-living α-proteobacteria (e.g., Andersson et al. 1998, 2003; Esser et al. 2004; Gray and Spencer 1996; Lang et al. 2005; Sicheritz-Ponten et al. 1998). Due to the high A + T content and high evolutionary rates of both mitochondria and *Rickettsia*-like bacteria, their adhesion may indeed be an artifact of phylogenetic inference. This issue is illustrated by a recent controversy over a group of marine bacteria (SAR11 clade) that was claimed to be related to mitochondria (Thrash et al. 2012). An artifactual attraction of the involved organismal groups (SAR11 bacteria share a high A + T gene content with Rickettsiales and mitochondria) was only detected when employing more sophisticated phylogenetic methods and more realistic evolutionary models (Brindefalk et al. 2011; Rodriguez-Ezpeleta and Embley 2012).

Given the genome sequence data available today, it is then indeed unfeasible either to demonstrate a single origin of mitochondria or to exclude more than one (in particular for species with hydrogenosomes or mitosomes with only few remaining mitochondrion-related genes). To resolve this issue, substantially more genomic data from *Rickettsia*-like bacteria is required, in particular from deeply diverging and slowly evolving ones to overcome potential phylogenetic artifacts. In addressing this question, it is also important to recognize that certain gene acquisitions may result from bacterial predation (following the provocative postulate "you are what you eat"; Doolittle 1998), and from transient bacteria–eukaryote symbiotic associations, adding to the challenges of phylogenetic inference and requiring an even better bacterial taxon sampling.

Finally, the investigation of present-time relationships of intracellular α-proteobacteria with eukaryotes may help evaluating the likelihood of a multiple mitochondrial origin. A great number of bacterial endosymbionts belong to the Rickettsiales, including the genera *Rickettsia*, *Anaplasma*, *Ehrlichia*, and *Wolbachia*. These are obligate, intracellular endosymbionts/parasites of eukaryotic (in most known instances animal) cells, superficially resembling pre-mitochondria. *Rickettsia*-like bacteria also reside in plants (Caspi-Fluger et al. 2011) and in protists (Horn et al. 1999), including the already mentioned *Holospora* and *Caedibacter* found in ciliates (Springer et al. 1993). In most of these examples, the bacterial partner depends on the eukaryotic host for propagation, whereas benefits to the host are at best circumstantial (e.g., Giorgini et al. 2010; Mercot

and Charlat 2004; Rigaud and Juchault 1993) or inexistent. In our studies, we have come across a remarkable case of mutual dependence of host and symbiont, in the heterolobosean amoeba *Stachyamoeba lipophora* (CCAP 1579/1). Attempts to sequence mtDNA from this species revealed a large (close to 2 Mb), A + T-rich DNA of a bacterial endosymbiont, which does not grow on any of the tested synthetic media. More recent genome sequencing identifies this bacterium as a deeply diverging, little derived member of Rickettsiaceae, with full capacity for oxidative phosphorylation and a seemingly complete set of genes for a flagellar apparatus (B. Franz Lang, unpublished). Its DNA occurs side by side with a canonical mtDNA that shares features with that of the heterolobosean amoeba *Naegleria*. While most other *Rickettsia*-like bacteria are transient guests in eukaryotes, the *Stachyamoeba* cells cannot be cured from the endosymbiont by antibiotic treatment. Reduction of the endosymbiont genome copy number occurs, but results in increasingly slower growth of the amoeba and ultimately in its death. The resemblance of this endosymbiont with an early proto-mitochondrial state is striking, suggesting the potential for a long-term transition to an organelle that may substitute the resident mitochondrion, should the latter become nonfunctional.

In conclusion, the *Stachyamoeba* example evokes the possibility of repeated mitochondriogenesis from within α-proteobacteria—and without invoking major benefits via cellular syntrophy. It seems imperative to revisit the view of a single mitochondrial origin by comprehensive phylogenomics.

The Archezoan (SET) Hypothesis on the Origin of Eukaryotes and Mitochondria

According to traditional views, eukaryotes evolved gradually from primitive organisms to more eukaryote-like complex cells (the amitochondriate archezoa), which at some time point in evolution acquired mitochondria through endosymbiosis: serial endosymbiosis theory (SET; Doolittle 1980, 1981; Margulis 1981; Taylor 1987). That extant eukaryotes are indeed genetic chimeras was confirmed early on by phylogenetic studies, in which mitochondria and plastids clearly group with two different bacterial lineages (Gray et al. 1984). In contrast, phylogenies with cytoplasmic ribosomal sequences supported three clearly distinct domains of life: Eukarya, Archaea, and Bacteria (Vossbrinck et al. 1987). The finding of unicellular, phagocytotic eukaryotes without mitochondria (initially termed archezoa; including Microsporidia, Metamonada such as *Giardia*, and Parabasalia such as *Trichomonas*) defined a further step in the search for eukaryotic origins (Cavalier-Smith 1983). These organisms live under anaerobic conditions and contain a regular nucleus, a cytoskeleton, and a more or less rudimentary endomembrane system. Initially, they were assumed to be primitively without mitochondria, implying that the hallmarks of eukaryotes were already present prior to endosymbiosis. This view was reinforced by some (but not all) ribosomal RNA (rRNA)-based phylogenies, placing archezoa as the earliest diverging eukaryotic lineages (e.g., Leipe et al. 1993; Sogin et al. 1989; Vossbrinck et al. 1987; Woese and Fox 1977).

The archezoan hypothesis came into discredit with two realizations. First, these rRNA-based tree topologies turned out to be plagued by phylogenetic artifacts, resulting in the attraction of fast-evolving species to a basal position (Budin and Philippe 1998; Kumar and Rzhetsky 1996; Philippe 2000; Philippe and Forterre 1999; Stiller and Hall 1999). Second, genes were discovered in these assumed ancestral organisms which are normally involved in mitochondrial biogenesis (such as heat shock proteins), and which in phylogenetic analyses group with mitochondrial/α-proteobacterial counterparts. Further, several of these proteins were shown to be targeted to either hydrogenosomes or previously unrecognized cryptic organelles (mitosomes or cryptons), which exist throughout all archezoan groups. This suggests that these organelles may all derive from mitochondria (e.g., Bui et al. 1996; Embley and Hirt 1998; Germot et al. 1996, 1997; Mai et al. 1999; Regoes et al. 2005; Roger et al. 1998; Tovar et al. 1999; Williams and Keeling 2003), although as already mentioned, an alternative α-proteobacterial source close to the mitochondrial origin is difficult to exclude.

In face of the above findings, it needs to be explained why all extant eukaryotes possess the distinctive characters of eukaryotes such as a nucleus and a cytoskeleton. In evolutionary terms, this favors the view that the host of the mitochondrial endosymbiont was a complex eukaryotic cell. But why then is there no trace of truly amitochondriate eukaryotes? The reason might be that biologists have failed to detect them or that these organisms became extinct, outcompeted by derived, mitochondriate, secondarily anaerobic eukaryotes. Curiously, this *lacking evidence* for primitively amitochondriate eukaryotes was proclaimed by some as "downfall of the archezoan hypothesis." Yet, lacking evidence for a given theory is different from evidence against it.

A Eukaryotic Origin by Symbiogenesis?

The analysis of complete eukaryotic genomes confirms that eukaryotes are most complex genetic mosaics, with a more archaea-related information processing machinery and bacteria-related operational genes involved in metabolism (Lopez-Garcia and Moreira 1999; Ribeiro and Golding 1998; Rivera et al. 1998; Zillig et al. 1989a, b). These findings have been rationalized by two lines of diametrically opposite interpretations on the origin of eukaryotes. One is that truly amitochondriate archezoa once existed and that these share an ancient evolutionary past with Archaea. Archezoa either became extinct or remain to be identified in strictly anoxic environments. The other interpretation is that eukaryotes emerged by "symbiogenesis" (Koonin 2010), involving metabolic symbiosis (syntrophy) between an archaeal and a bacterial partner (Fig. 1). The probably best known theory of this genre is the hydrogen hypothesis (Martin and Müller 1998), which posits that an α-proteobacterium with the capacity of oxidative phosphorylation played the role of the symbiotic bacterial partner that produced hydrogen under anaerobic, heterotrophic conditions and that the hydrogen was utilized by a hydrogen-dependent, autotrophic archaebacterial host. Eventually, the α-proteobacterium would transform into the semiautonomous mitochondrial organelle. The integration of both metabolic systems would ultimately give rise to the

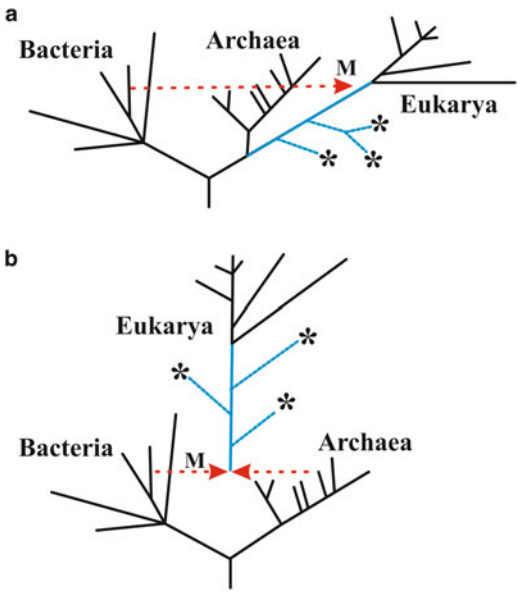


Fig. 1 Alternative hypotheses on the origin of the eukaryotic cell. (**a**) The iconic three-domain tree of life based on rRNA sequence data [tree redrawn from Woese et al. (1990)], assuming an origin of eukaryotes by stepwise evolution from an amitochondriate eukaryote (archezoan; *blue line*, extinct *blue* lineages marked with *asterisks*) and subsequent acquisition of the mitochondrion (M) through endosymbiosis with an α-proteobacterium. Note that the rooting of the tree is arbitrarily chosen within bacteria (see also text). (**b**) Symbiogenesis scenario (hydrogen hypothesis), redrawn from (**a**) with the same phylogenetic distances. Note that eukaryotic evolution has to be extremely rapid, when compared to (**a**). This interpretation implies a major extinction period (*blue line*; extinct *blue* lineages marked with *asterisks*) prior to the divergence of extant eukaryotes. The tree shape is inconsistent with the mitochondrial "big-bang"-shaped phylogenies that rather fit scenario (**a**)

typical eukaryotic metabolism, with one portion of it remaining localized in mitochondria. Similar hypotheses assume other and additional bacterial partners and a different timing when mitochondria come into play (e.g., Forterre 2010; Moreira and Lopez-Garcia 1998), but phylogenetic evidence for these more complicated scenarios is weak. In the following we will focus on the theoretical implications of the hydrogen hypothesis and the underlying symbiogenesis concept. For more details, see Cavalier-Smith (2009), Embley and Martin (2006), Forterre (2010), Gray and Archibald (2012), Koonin (2010), and Poole and Penny (2007).

Implications of the Hydrogen Hypothesis

According to the hydrogen hypothesis for the origin of eukaryotes, uptake of the α-proteobacterium that gives rise to mitochondria triggers an accelerated evolution of the archaeal host, leading to the nucleus, cytoskeleton, cellular compartmentalization,

the cellular machinery for microbial predation, etc. When placed into an evolutionary context, this postulate implies several inconsistencies, some of which also apply to other flavors of symbiogenesis hypotheses.

1. *Mitochondriate cells without typical eukaryotic features?* Following the hydrogen hypothesis, transitional cell forms must have once existed that had a mitochondrion, but no nucleus, cytoskeleton, and endomembrane system. Descendents of such organisms are unknown. While the classic archezoan scenario struggles with lacking evidence for truly *amitochondriate* eukaryotes, the hydrogen hypothesis struggles with lacking evidence for intermediate forms of *mitochondriate* eukaryotes. It is difficult to rationalize that mitochondriate precursors of eukaryotes disappeared, despite an alleged advantage through metabolic syntrophy and tolerance of oxygen in an increasingly oxygen-rich environment.
2. *Archaeal methanogens close to eukaryotes?* Most symbiogenesis theories postulate an archaebacterial methanogen as a host. If that were true, archaea-related eukaryotic genes would resemble genes of extant methanogens in sequence comparisons, with similarly high scores as genes of mitochondrial origin with α-proteobacterial genes. Since this is not observed, the only way out is that this entire line of methanogens became extinct.
3. *Clear phylogenetic trace of the α-proteobacterial, but not for the archaeal partner*. If eukaryotes indeed derived via symbiogenesis of an archaeal methanogen with an α-proteobacterium, phylogenies with eukaryotic genes of mitochondrial versus archaebacterial origin should identify the same alleged fusion time point. In addition, genes of the archaeal host should have undergone less genetic change than those of the symbiont, because genes transferred to the host's nucleus have to be adapted to new rules of gene regulation and expression, and for gene products functioning in the mitochondrion, targeting signals have to be invented for import into the organelle. Examples of currently ongoing mitochondrial gene transfer demonstrate that substantial sequence changes occur (Adams et al. 1999; Archibald and Richards 2010). Therefore, when comparing phylogenies with genes of mitochondrial and archaeal origin, the latter should be characterized by shorter branch length and higher phylogenetic resolution. Yet, quite the opposite is observed: mitochondrial phylogenies, whether with mtDNA-encoded genes or those transferred to the nucleus, clearly point to a mitochondrial origin within proteobacteria or even within α-proteobacteria. In addition, despite their accelerated evolution, mitochondrial genes allow rather confident rooting of the eukaryotic tree (Derelle and Lang 2011). In contrast, even aligning the most conserved archaea-related eukaryotic protein sequences with their archaeal counterparts is difficult, for half or more of sequence positions (Cox et al. 2008). This cannot be explained by accelerated evolution of archaea-related genes in the eukaryotic nucleus, because most mitochondrial (those that were transferred in an early stage of symbiosis) and archaeal genes in the nucleus would coevolve, and if acceleration took place it would apply only to those of mitochondrial origin as explained above.
4. *Molecular machines are distinct and eukaryote-specific in the cytoplasm—but bacteria-like in mitochondria*. The three-domain classification of organismal life is not just a phylogenetic invention based on rRNA sequence data, but makes reference

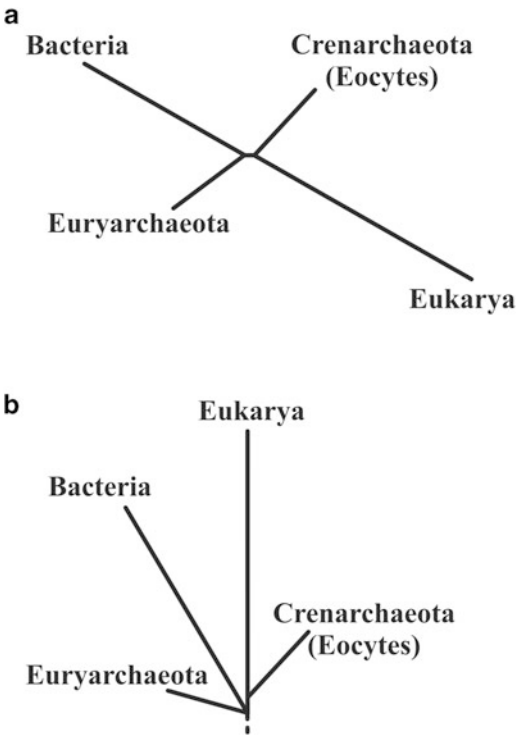


Fig. 2 Phylogenetic tree with 49 archaea-related eukaryotic proteins. Schematically redrawn, using inferences published by Cox et al. (2008). (**a**) Unrooted tree, with four rather than three separate phylogenetic domains. (**b**) The same tree, rooted arbitrarily at the divergence of Euryarchaeota and Bacteria

to three structurally and functionally distinct ribosomal archetypes. As correctly pointed out (Forterre 2010), the eukaryotic ribosome is so distinct that it is difficult to argue that it was introduced by a fusion with an archaeal partner. For cytoplasmic ribosomes to be that archetypical, one had to postulate either a most ancient fusion event leaving ample time for evolutionary change (i.e., either assume a very ancient origin of mitochondria, which is inconsistent with phylogenetic inference as explained above, or postulate an ancient symbiogenesis event that differs from the mitochondrial one) or extreme evolutionary acceleration. However, mitochondrial ribosomes in many protist lineages (e.g., the jakobid *Reclinomonas americana*; Lang et al. 1997) are clearly of the bacterial type, and this despite the relative fast evolution of bacteria, and regardless of mitochondria undergoing reductive, highly accelerated evolution compared to bacteria. Clearly, there is no evidence for highly accelerated evolution of cytoplasmic eukaryotic ribosomes; the typical structural features must be the product of a long evolutionary past that is proper to Eukarya. The same arguments are similarly valid for RNA polymerases (Lang et al. 1997) and RNase P RNAs in mitochondria of jakobids (Seif et al. 2006) that are strikingly close to bacterial counterparts, but clearly distinct from cytoplasmic forms.

5. *Archaea-related eukaryotic genes do not derive from methanogens*. Phylogenies including archaea-related eukaryotic genes and genes from extant Archaea do not group within methanogens but instead with eocytes (Fig. 2), with high statistical confidence (Cox et al. 2008). Is it possible that the predicted tree

topology is due to a phylogenetic long-branch-attraction (LBA) artifact? This is unlikely. When evoking LBA for this dataset, one might instead expect attraction of the two fast-evolving groups: Eukarya and Bacteria.

6. *Distinct class of well-conserved eukaryote-specific genes in the nucleus.* When selecting genes for phylogenomic analyses, one is confronted with at least four major categories of eukaryotic genes. The largest fraction is fast-evolving genes with sometimes obscure evolutionary path including laterally transferred genes, potential sources of δ-proteobacterial genes (Moreira and Lopez-Garcia 1998), chlamydial genes that are somehow connected to photosynthesis (Huang and Gogarten 2007), and potential viral sources (Forterre 2010; Richards and Archibald 2011). In one publication (Ribeiro and Golding 1998), a large fraction of eukaryotic genes is therefore (cautiously) qualified as of Gram-negative bacterial origin. Most of these genes retain too little phylogenetic signal to reveal a more precise origin or to be useful in phylogenomic inferences.

   The remaining well-conserved eukaryotic genes divide into three about equal-sized groups: archaea-related, proteobacteria-related, and . . . eukaryote-specific ones coding for functions that set eukaryotes apart from the rest. In discussing the origin of eukaryotes, this latter group of enigmatic genes is central. It is unreasonable to postulate that such a large group of *well-conserved* genes was created within a record short period right after an alleged fusion event and without a reliable trace of origin (and without postulating the same for archaea-related and proteobacteria-related genes). Rather, these enigmatic eukaryote-specific genes can be expected to having evolved over a long period of evolutionary time, in a line of now extinct eukaryotic ancestors. In the archezoan scenarios this period would match truly amitochondriate eukaryotes, i.e., before the introduction of mitochondria. In case of the hydrogen hypothesis, mitochondria would have to be introduced early to allow for evolution of this group of eukaryote-specific genes. Yet, this is in conflict not only with phylogenetic arguments raised in the previous points. A mitochondrion-early scenario implies that oxidative phosphorylation was introduced at a time point when oxygen concentrations were low in the atmosphere and that the full gene set was preserved over a very long evolutionary period. This seems unlikely when looking into present-time anoxic habitats. These are populated with eukaryotes that have hydrogenosomes or mitosomes, no oxidative phosphorylation as it is of no use, and some level of oxygen tolerance providing a competitive edge over strictly anaerobic organisms.
7. *How did phagocytosis and endosymbiosis evolve?* From a cell biology view, extant eukaryotes stand out by complex cellular structures and machineries that allow for both phagocytosis and endosymbiosis, features that are virtually nonexistent in Bacteria and Archaea. The host cell engulfing the α-proteobacterial precursor of mitochondria likely possessed these capacities—a description that fits that of an archezoan [an idea emphasized by Cavalier-Smith (2009), Poole and Neumann (2010), and Poole and Penny (2007); although restricted to phagocytosis]. The seemingly most rapid, *big-bang-like* radiation of eukaryotes (Philippe et al. 2000) into diverse lineages reinforces this idea (Brinkmann and Philippe 2007). It seems that following introduction of mitochondria no time was spent on evolving complex eukaryote-specific structures, and on inventing the many hundreds of

eukaryote-specific genes that encode components of the nucleus, flagellar apparatus, etc.—that are unaccounted for by symbiogenesis scenarios. The finding of a few eukaryote-related genes in Archaea that do not exist in bacteria does not help to change this view. For instance, the histone-like genes in both major groups of Archaea (Cubonova et al. 2005; Sandman and Reeve 2005) may be interpreted as genuine precursors of eukaryotic histones, but also as simplified archaeal versions that stem from ancient Eukarya (e.g., Philippe and Forterre 1999), and the fast-evolving bacteria might have once had histones that were lost at some stage of reductive evolution. In the end, the principle of parsimonious gene counting used out of phylogenetic context is of little value for inferring evolutionary scenarios.

8. *Evolution proceeds from simple to complex?* Symbiogenesis theories assume a priori that complex eukaryotic cells can only derive from simple prokaryotic ones. Yet, there is no evidence that evolution always follows a path from simple (primitive, low) to complex, a popular concept borrowed from homocentric theories. It rather seems that the success of eukaryotic lineages is due to simplification and streamlining, following symbioses that introduce the challenge of managing multiple genomes, and most different replication, transcription, and translation machineries whose activities need to be coordinated. The vanishing endosymbiont and organelle genomes witness this evolutionary trend, so do eukaryotes adapted to anaerobic environments that are living examples of reductive evolution, so are yeasts that are among the most successful eukaryotes despite a highly reduced nuclear genome.

Conclusions

At present, we have solid phylogenetic evidence for a relatively recent origin of mitochondria, which allows rooting of extant mitochondriate eukaryotes, and which will allow more precise identification of the mitochondrial origin within Proteobacteria. We have further strong phylogenetic evidence that a portion of eukaryotic genes relates to Crenarchaeota (although this does not mean that eukaryotes originated from this archaeal group; the opposite may also be true). This raises the possibility of four instead of three domains of life: Euryarchaeota, Bacteria, Crenarchaeota, and Eukarya (Fig. 2). However, the question where to root the tree of life remains unanswered. The widespread assumption that places the root in prokaryotes is unfounded, as phylogenies based on gene paralogs are based on only few informative sequence sites and therefore unreliable. On the other hand, an origin of cellular life from within eukaryotes remains an option, since there are many examples demonstrating the evolution of structurally and genetically simple from complex organisms.

All current hypotheses on the origin of eukaryotes assume extinction of transitory lineages, one event for the archezoan theory and several for the hydrogen hypothesis. With virtually all permutations of theories on the origin of eukaryotes postulated, it now seems timely to go back to work: search for missing descendants

of predicted eukaryotic precursors, sequence and analyze genomes from a much broader collection of protists and bacteria close to the mitochondrial origin, refine phylogenomic analyses, and improve evolutionary models and inferences to avoid systematic error.

Acknowledgments The author is grateful to Dr. G. Burger (Université de Montréal) for comments on the manuscript and to NSERC and Genome Quebec/Genome Canada for providing financial support.

Glossary

Archaea Organisms constituting one domain of the iconic three-domains-of-life hypothesis; subdivided into Crenarchaeota (Eocytes) and Euryarchaeota (including methanogens).

Archezoa Hypothetical eukaryotes that never had a mitochondrion; not to be confused with amitochondriate species that secondarily lost mitochondria and typically carry vestiges of this organelle.

CAT One of the models used in phylogenetic reconstruction based on protein sequence; uses categories of distinct, site-wise amino acid profiles (inferred from multiple sequence alignment); inferences with this model are arguably least prone to LBA.

Eocyte One of the two large archaeal clades; according to phylogenetic analyses, the main source of archaea-related genes in extant eukaryotes.

Informational genes Genes involved in genetic information transfer and processing; principal components of replication, transcription, and translation.

LBA Long-Branch Attraction; phylogenetic artifact that leads to the incorrect grouping of fast-evolving species, or attraction to a distant outgroup, due to evolutionary model violations and underestimation of repeated sequence changes.

mtDNA Mitochondrial DNA.

Operational genes Genes involved in biosynthesis and metabolism.

Phagocytosis Uptake of a cell by another cell, usually by sequestration in a food vacuole.

Phylogenomics Phylogenetic inferences based on genome-wide selections of genes that have to be sufficiently conserved to contribute to a phylogenetic signal (i.e., to avoid introduction of "noise" and "sequence bias"); avoiding paralogs and potential laterally transferred genes that are inconsistent with the species tree.

Protists Eukaryotes other than fungi, animals, and plants; predominantly unicellular. Protists represent more biological diversity than the three former groups together. Fungi, animals, and plants emerged from within protists.

Symbiogenesis A concept claiming endosymbiosis as the reason for highly accelerated evolution of new eukaryotic features. Symbiogenesis theories have been also termed "fusion theories," in which "fusion" stands for endosymbiosis in a wide sense.

References

Adams KL, Song K, Roessler PG, Nugent JM, Doyle JL et al (1999) Intracellular gene transfer in action: dual transcription and multiple silencings of nuclear and mitochondrial *cox2* genes in legumes. Proc Natl Acad Sci USA 96:13863–13868

Andersson SG, Zomorodipour A, Andersson JO, Sicheritz-Ponten T, Alsmark UC et al (1998) The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. Nature 396:133–140

Andersson SG, Karlberg O, Canback B, Kurland CG (2003) On the origin of mitochondria: a genomics perspective. Philos Trans R Soc Lond B Biol Sci 358:165–177, discussion 77–9

Archibald JM, Richards TA (2010) Gene transfer: anything goes in plant mitochondria. BMC Biol 8:147

Brindefalk B, Ettema TJ, Viklund J, Thollesson M, Andersson SG (2011) A phylometagenomic exploration of oceanic alphaproteobacteria reveals mitochondrial relatives unrelated to the SAR11 clade. PLoS One 6:e24457

Brinkmann H, Philippe H (2007) The diversity of eukaryotes and the root of the eukaryotic tree. Adv Exp Med Biol 607:20–37

Budin K, Philippe H (1998) New insights into the phylogeny of eukaryotes based on ciliate Hsp70 sequences. Mol Biol Evol 15:943–956

Bui ET, Bradley PJ, Johnson PJ (1996) A common evolutionary origin for mitochondria and hydrogenosomes. Proc Natl Acad Sci USA 93:9651–9656

Caspi-Fluger A, Inbar M, Mozes-Daube N, Katzir N, Portnoy V et al (2011) Horizontal transmission of the insect symbiont *Rickettsia* is plant-mediated. Proc R Soc B 279:1791–1796

Cavalier-Smith T (1983) Members of the Archezoa include the Microsporidia, Metamonada, and Parabasalia. In: Schwemmler W, Schenk HEA (eds) Endocytobiology II. De Gruyter, Berlin, pp 1027–1034

Cavalier-Smith T (2009) Predation and eukaryote cell origins: a coevolutionary perspective. Int J Biochem Cell Biol 41:307–322

Cox CJ, Foster PG, Hirt RP, Harris SR, Embley TM (2008) The archaebacterial origin of eukaryotes. Proc Natl Acad Sci USA 105:20356–20361

Cubonova L, Sandman K, Hallam SJ, Delong EF, Reeve JN (2005) Histones in crenarchaea. J Bacteriol 187:5482–5485

Derelle R, Lang BF (2011) Rooting the eukaryotic tree with mitochondrial and bacterial proteins. Mol Biol Evol 29:1277–1289

Doolittle WF (1980) Revolutionary concepts in evolutionary cell biology. Trends Biochem Sci 5:146–149

Doolittle WF (1981) The endosymbiont hypothesis. Science 213:640–641

Doolittle WF (1998) You are what you eat: a gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. Trends Genet 14:307–311

Embley TM, Hirt RP (1998) Early branching eukaryotes? Curr Opin Genet Dev 8:624–629

Embley TM, Martin W (2006) Eukaryotic evolution, changes and challenges. Nature 440:623–630

Esser C, Ahmadinejad N, Wiegand C, Rotte C, Sebastiani F et al (2004) A genome phylogeny for mitochondria among alpha-proteobacteria and a predominantly eubacterial ancestry of yeast nuclear genes. Mol Biol Evol 21:1643–1660

Forterre P (2010) A new fusion hypothesis for the origin of Eukarya: better than previous ones, but probably also wrong. Res Microbiol 162:77–91
Fujishima M, Fujita M (1985) Infection and maintenance of *Holospora obtusa*, a macronucleus-specific bacterium of the ciliate *Paramecium caudatum*. J Cell Sci 76:179–187
Fujishima M, Dohra H, Kawai M (1997) Quantitative changes in periplasmic proteins of the macronucleus-specific bacterium *Holospora obtusa* in the infection process of the ciliate *Paramecium caudatum*. J Eukaryot Microbiol 44:636–642
Germot A, Philippe H, Le Guyader H (1996) Presence of a mitochondrial-type 70-kDa heat shock protein in *Trichomonas vaginalis* suggests a very early mitochondrial endosymbiosis in eukaryotes. Proc Natl Acad Sci USA 93:14614–14617
Germot A, Philippe H, Le Guyader H (1997) Evidence for loss of mitochondria in Microsporidia from a mitochondrial-type HSP70 in *Nosema locustae*. Mol Biochem Parasitol 87:159–168
Giorgini M, Bernardo U, Monti MM, Nappo AG, Gebiola M (2010) *Rickettsia* symbionts cause parthenogenetic reproduction in the parasitoid wasp *Pnigalio soemius* (Hymenoptera: Eulophidae). Appl Environ Microbiol 76:2589–2599
Görtz HD, Lellig S, Miosga O, Wiemann M (1990) Changes in fine structure and polypeptide pattern during development of *Holospora obtusa*, a bacterium infecting the macronucleus of *Paramecium caudatum*. J Bacteriol 172:5664–5669
Gray MW, Archibald JM (2012) Genomics of chloroplasts and mitochondria. In: Bock R, Knoop V (eds) Origins of mitochondria and plastids. Springer, Dordrecht
Gray M, Spencer D (1996) Organellar evolution. In: Roberts D, Sharp P, Alderson G, Collins M (eds) Evolution of microbial life. Cambridge University Press, Cambridge, pp 109–26
Gray MW, Sankoff D, Cedergren RJ (1984) On the evolutionary descent of organisms and organelles: a global phylogeny based on a highly conserved structural core in small subunit ribosomal RNA. Nucleic Acids Res 12:5837–5852
Horn M, Fritsche TR, Gautom RK, Schleifer KH, Wagner M (1999) Novel bacterial endosymbionts of *Acanthamoeba spp*. related to the *Paramecium caudatum* symbiont *Caedibacter caryophilus*. Environ Microbiol 1:357–367
Huang J, Gogarten JP (2007) Did an ancient chlamydial endosymbiosis facilitate the establishment of primary plastids? Genome Biol 8:R99
Imanian B, Pombert JF, Keeling PJ (2010) The complete plastid genomes of the two 'dinotoms' *Durinskia baltica* and *Kryptoperidinium foliaceum*. PLoS One 5:e10711
Koonin EV (2010) The origin and early evolution of eukaryotes in the light of phylogenomics. Genome Biol 11:209
Kumar S, Rzhetsky A (1996) Evolutionary relationships of eukaryotic kingdoms. J Mol Evol 42:183–193
Lang BF, Burger G, O'Kelly CJ, Cedergren R, Golding GB et al (1997) An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. Nature 387:493–497
Lang BF, Brinkmann H, Koski L, Fujishima M, Goertz HD, Burger G (2005) On the origin of mitochondria and *Rickettsia*-related eukaryotic endosymbionts. Jpn J Protozool 38:171–183
Leipe DD, Gunderson JH, Nerad TA, Sogin ML (1993) Small subunit ribosomal RNA + of *Hexamita inflata* and the quest for the first branch in the eukaryotic tree. Mol Biochem Parasitol 59:41–48
Lopez-Garcia P, Moreira D (1999) Metabolic symbiosis at the origin of eukaryotes. Trends Biochem Sci 24:88–93
Mai Z, Ghosh S, Frisardi M, Rosenthal B, Rogers R, Samuelson J (1999) Hsp60 is targeted to a cryptic mitochondrion-derived organelle ("crypton") in the microaerophilic protozoan parasite *Entamoeba histolytica*. Mol Cell Biol 19:2198–2205
Margulis L (1981) Symbiosis in cell evolution. Freeman, San Francisco, CA
Martin W, Müller M (1998) The hydrogen hypothesis for the first eukaryote. Nature 392:37–41
Mercot H, Charlat S (2004) *Wolbachia* infections in *Drosophila melanogaster* and *D. simulans*: polymorphism and levels of cytoplasmic incompatibility. Genetica 120:51–59

Moreira D, Lopez-Garcia P (1998) Symbiosis between methanogenic archaea and delta-proteobacteria as the origin of eukaryotes: the syntrophic hypothesis. J Mol Evol 47:517–530
Philippe H (2000) Opinion: long branch attraction and protist phylogeny. Protist 151:307–316
Philippe H, Forterre P (1999) The rooting of the universal tree of life is not reliable. J Mol Evol 49:509–523
Philippe H, Lopez P, Brinkmann H, Budin K, Germot A et al (2000) Early-branching or fast-evolving eukaryotes? An answer based on slowly evolving positions. Proc R Soc Lond B Biol Sci 267:1213–1221
Poole AM, Neumann N (2010) Reconciling an archaeal origin of eukaryotes with engulfment: a biologically plausible update of the Eocyte hypothesis. Res Microbiol 162:71–76
Poole AM, Penny D (2007) Evaluating hypotheses for the origin of eukaryotes. Bioessays 29:74–84
Regoes A, Zourmpanou D, Leon-Avila G, van der Giezen M, Tovar J, Hehl AB (2005) Protein import, replication, and inheritance of a vestigial mitochondrion. J Biol Chem 280:30557–30563
Ribeiro S, Golding GB (1998) The mosaic nature of the eukaryotic nucleus. Mol Biol Evol 15:779–788
Richards TA, Archibald JM (2011) Cell evolution: gene transfer agents and the origin of mitochondria. Curr Biol 21:R112–R114
Rigaud T, Juchault P (1993) Conflict between feminizing sex ratio distorters and an autosomal masculinizing gene in the terrestrial isopod *Armadillidium vulgare* Latr. Genetics 133:247–252
Rivera MC, Jain R, Moore JE, Lake JA (1998) Genomic evidence for two functionally distinct gene classes. Proc Natl Acad Sci USA 95:6239–6244
Rodriguez-Ezpeleta N, Embley TM (2012) The SAR11 group of alpha-proteobacteria is not related to the origin of mitochondria. PLoS One 7:e30520
Roger AJ, Svard SG, Tovar J, Clark CG, Smith MW et al (1998) A mitochondrial-like chaperonin 60 gene in *Giardia lamblia*: evidence that diplomonads once harbored an endosymbiont related to the progenitor of mitochondria. Proc Natl Acad Sci USA 95:229–234
Sandman K, Reeve JN (2005) Archaeal chromatin proteins: different structures but common function? Curr Opin Microbiol 8:656–661
Seif E, Cadieux A, Lang BF (2006) Hybrid *E. coli* – mitochondrial ribonuclease P RNAs are catalytically active. RNA 12:1661–1670
Sicheritz-Ponten T, Kurland CG, Andersson SG (1998) A phylogenetic analysis of the cytochrome b and cytochrome c oxidase I genes supports an origin of mitochondria from within the Rickettsiaceae. Biochim Biophys Acta 1365:545–551
Sogin ML, Gunderson JH, Elwood HJ, Alonso RA, Peattie DA (1989) Phylogenetic meaning of the kingdom concept: an unusual ribosomal RNA from *Giardia lamblia*. Science 243:75–77
Springer N, Ludwig W, Amann R, Schmidt HJ, Görtz HD, Schleifer KH (1993) Occurrence of fragmented 16S rRNA in an obligate bacterial endosymbiont of *Paramecium caudatum*. Proc Natl Acad Sci USA 90:9892–9895
Stiller JW, Hall BD (1999) Long-branch attraction and the rDNA model of early eukaryotic evolution. Mol Biol Evol 16:1270–1279
Taylor FJ (1987) An overview of the status of evolutionary cell symbiosis theories. Ann N Y Acad Sci 503:1–16
Thrash JC, Boyd A, Huggett MJ, Grote J, Carini P et al (2012) Phylogenomic evidence for a common ancestor of mitochondria and the SAR11 clade. Sci Rep 1:13
Tovar J, Fischer A, Clark CG (1999) The mitosome, a novel organelle related to mitochondria in the amitochondrial parasite *Entamoeba histolytica*. Mol Microbiol 32:1013–1021
Vossbrinck CR, Maddox JV, Friedman S, Debrunner-Vossbrinck BA, Woese CR (1987) Ribosomal RNA sequence suggests microsporidia are extremely ancient eukaryotes. Nature 326:411–414
Wiemann M, Görtz HD (1991) Identification and localization of major stage-specific polypeptides of infectious *Holospora obtusa* with monoclonal antibodies. J Bacteriol 173:4842–4850

Williams BA, Keeling PJ (2003) Cryptic organelles in parasitic protists and fungi. Adv Parasitol 54:9–68

Wixon J (2001) Featured organism: reductive evolution in bacteria: *Buchnera* sp., *Rickettsia prowazekii* and *Mycobacterium leprae*. Comp Funct Genomics 2:44–48

Woese CR, Fox GE (1977) Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc Natl Acad Sci USA 74:5088–5090

Woese CR, Kandler O, Wheelis ML (1990) Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc Natl Acad Sci USA 87:4576–4579

Zillig W, Klenk H-P, Palm P, Leffers H, Pühler G et al (1989a) Did eukaryotes originate by a fusion event? Endocytobiosis Cell Res 6:1–25

Zillig W, Klenk HP, Palm P, Puhler G, Gropp F et al (1989b) The phylogenetic relations of DNA-dependent RNA polymerases of archaebacteria, eukaryotes, and eubacteria. Can J Microbiol 35:73–80

Modifications and Innovations in the Evolution of Mitochondrial Protein Import Pathways

Victoria Hewitt, Trevor Lithgow, and Ross F. Waller

Abstract Eukaryotic cells are defined by their mitochondria, organelles that were derived through endosymbiosis. The development of this organelle from a bacterial endosymbiont required establishment of effective protein import pathways so that much of the genetic capacity of the bacterium could be relocated to the host cell. Two realms of study have delivered insight into the early evolution of these mitochondrial pathways: (1) considering the "starting material" based on what can be observed of protein trafficking pathways in extant species of bacteria and (2) analysing the protein import pathways of parasites whose mitochondria have undergone secondary reduction and now offer insight into minimal functional pathways. These approaches have illuminated what components of bacterial trafficking pathways were co-opted in the developing mitochondrion and what further innovations occurred within the eukaryote host. Now comparative analysis of model mitochondrial systems, with organelles found in a broad diversity of eukaryotes (namely protists), shows when in eukaryotic radiation these major innovations took place and what lineage-specific changes have since occurred to mitochondrial import systems in eukaryotes.

V. Hewitt
Department of Biochemistry and Molecular Biology, Monash University, Clayton Campus, Melbourne 3800, Australia

T. Lithgow (✉)
Department of Microbiology, Monash University, Clayton Campus, Melbourne 3800, Australia
e-mail: trevor.lithgow@monash.edu

R.F. Waller (✉)
School of Botany, University of Melbourne, Melbourne 3010, Australia
e-mail: r.waller@unimelb.edu.au

W. Löffelhardt (ed.), *Endosymbiosis*, DOI 10.1007/978-3-7091-1303-5_2,
© Springer-Verlag Wien 2014

Contents

Introduction 20
An Ancestral System Doing (More or Less) What It Has Always Done 21
Ancestral Machines Have Been Modified and Recombined in the Course of Evolution 23
New Machines—Without a (Bacterial) Trace 25
New Insights into Organelle Evolution from Mitochondrial Diversity 26
References 30

Introduction

Like metabolic and signal transduction pathways, protein transport pathways have been revealed component by component through diverse experimental approaches and using diverse model organisms. In many cases, including the protein import pathway into mitochondria, we now have sufficient understanding to address holistic questions about protein transport pathways, such as how transport along a pathway is regulated, how the flux of a protein transport pathway impacts on other (metabolic, signal transduction, gene regulatory) pathways and, the subject of this review, how the pathways evolved. The conversion of an ancestral endosymbiont to a mitochondrion involved the transfer of genes from the bacterial population to the host cell genome (Andersson and Kurland 1999; Cavalier-Smith 2002, 2006; de Duve 2007; Embley and Martin 2006; Gray et al. 1999, 2001; Lang et al. 1999; Timmis et al. 2004). For this process to take effect the gene products translated in the cytosol had to be recognised and imported into mitochondria. The changes to the endosymbiont in its conversion to an organelle could not have occurred simultaneously and, accordingly, our understanding of the evolution of protein import pathways is developing through the characterisation of key events in this stepwise evolutionary process.

Mitochondrial protein import depends on the activity of a series of molecular machines. There are three themes that emerge when investigating the evolution of the mitochondrial protein transport machinery. Firstly, some pre-existing protein transport machines derived from bacteria were used directly, maintaining their ancestral function in mitochondria. Secondly, some parts of the bacterial machinery have been modified and recombined so that their ancestral biochemical function is adapted to perform new functions in a protein transport pathway. Thirdly, some machines bear no sequence similarity to bacterial proteins suggesting a non-bacterial origin, and examples from parasites give clues as to where and how these factors arose. Here we review how the endosymbiont's own protein transport machinery was adapted to new roles, how entirely novel machinery arose from pre-existing component parts and how embellishments unique to this particular host–symbiont relationship were set in place in the evolution of mitochondria. We will also consider how an increasingly expanded view of mitochondrial diversity in eukaryotes is shaping our understanding of the development of these pathways.

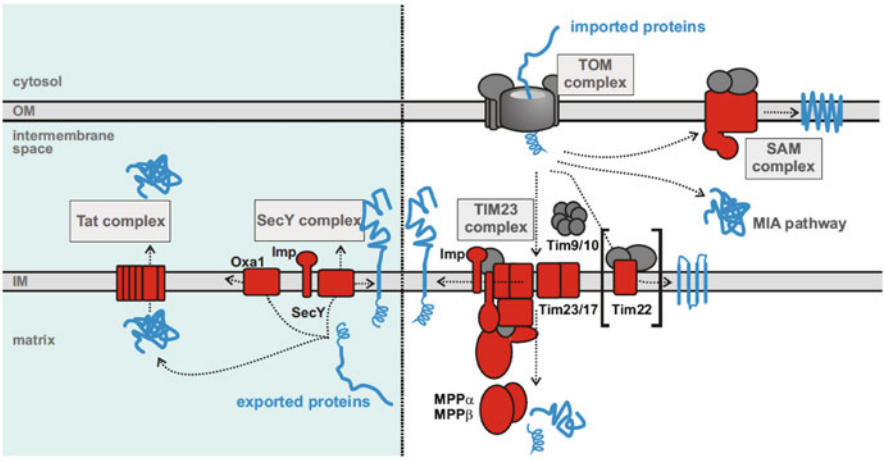


Fig. 1 Ancestral and derived import machines used in extant mitochondria. Membrane-located translocation machines are positioned in the outer and inner mitochondrial membranes, where they effect protein import into the organelle and assembly of integral proteins into the mitochondrial membranes. The details of the components of these machines have been recently reviewed (Chacinska et al. 2009; Neupert and Herrmann 2007). The *blue panel* summarises the machinery required for assembly of matrix-encoded membrane proteins into the inner membrane. *Red shading* denotes proteins for which bacterial proteins have been identified as ancestral; *grey shading* denotes components for which no bacterial ancestor has been identified

An Ancestral System Doing (More or Less) What It Has Always Done

The major biosynthetic route for bacterial membrane and periplasmic (intermembrane space) proteins is outward, from their synthetic origin in the cytoplasm to their peripheral destination. The bacterial protein trafficking pathways conserved in mitochondria, therefore, reflect this outward orientation. Initially these pathways would have supported only organelle-encoded gene products. Yet as more genes moved to the nucleus, some of these proteins would come to journey first across the mitochondrial membranes (either one or two) and then be redirected into their correct destination by these relic bacterial systems. Despite differences in nomenclature, four such "bacterial" systems are remnant in mitochondria—SecYEG, YidC (Oxa1), BAM (SAM) and Tat (discussed in section "New Insights into Organelle Evolution from Mitochondrial Diversity").

SecYEG and YidC translocases function in protein transport across bacterial inner membranes, are universally found in bacteria and therefore would have been present in the ancestral endosymbiont to mediate protein assembly into the inner membrane. The mitochondrial OXA (Oxidase Assembly) translocase was derived from the endosymbiont's YidC (Kiefer and Kuhn 2007; Ott and Herrmann 2010; Pohlschroder et al. 2005; Stuart 2002; van der Laan et al. 2005) and fulfils a functionally homologous role in mitochondria (Fig. 1). The mitochondrial core

subunit Oxa1 can complement the lack of yidC2 from *Streptococcus mutans*; YidC2 from this bacterium can also function in place of Oxa1 in *Saccharomyces cerevisiae* (Funes et al. 2009). The Oxa1 protein is widely found in eukaryotes (Dolezal et al. 2006), though it has been lost from organisms such as *Cryptosporidium parvum* that lack a mitochondrial genome and therefore do not assemble membrane proteins from the internal face of the inner membrane (Alcock et al. 2012). The SecYEG complex can function alone to transfer proteins into the bacterial periplasm (equivalent to the mitochondrial intermembrane space) and to insert membrane proteins with simple topologies into the inner membrane (Driessen et al. 2001) (Fig. 1). The SecYEG complex can also function in concert with YidC to assemble membrane proteins with more complicated topologies (Samuelson et al. 2000). SecYEG would have been present in the early protomitochondria: even today some jakobid protists, such as *Reclinomonas americana*, encode a SecY translocase in their mitochondrial DNA [see section "New Machines—Without a (Bacterial) Trace" for discussion] (Lang et al. 1997; Tong et al. 2011). In most mitochondria the SecYEG has been lost, but strikingly, for proteins of complex membrane topology, the inner membrane TIM complex cooperates with the OXA complex just as the bacterial SecYEG and YidC complexes do (van der Laan et al. 2005; Webb and Lithgow 2010).

The bacterial origin of the SAM (Sorting and Assembly Machinery) complex in the mitochondrial outer membrane is even clearer; it evolved from the BAM complex, found in the outer membrane of all Gram-negative bacteria (Fig. 1) (Cavalier-Smith 2006; Dolezal et al. 2006; Gentle et al. 2004; Gross and Bhattacharya 2009). The SAM and the BAM complexes fulfil the same function: inserting beta-barrels from the internal face of the outer membrane, although in the case of mitochondria these proteins are first imported across this outer membrane. The core components of the SAM and BAM complexes have a common ancestry and are known as Sam50 in mitochondria and BamA (also known as Omp85) in bacteria (Gatsos et al. 2008; Gentle et al. 2004, 2005; Knowles et al. 2009; Ruiz et al. 2006). While the SAM and BAM complexes are functionally homologous, significant evolution is evident in the mitochondrial SAM complex. Mitochondria have lost whole aspects of bacterial envelope biogenesis including the ability to synthesise lipoproteins (Gabaldon and Huynen 2007) determining the loss of the lipoprotein partners: BamB, BamC, BamD and BamE from the endosymbiont's BAM complex (Gatsos et al. 2008; Knowles et al. 2009; Ruiz et al. 2006). These have been replaced, either during or subsequent to this period of lipoprotein loss, by key proteins of uncertain ancestry. From functional studies we know that the metaxins, and perhaps other proteins too, serve as "modules" of the SAM complex (discussed further below). Further analysis of these components promises a fuller understanding of how the mitochondrial SAM complex evolved. It is reasonable to imagine that there was a relatively seamless transition between the two forms of the essential Omp85 protein at the core of the complex (BamA/Sam50). Further mechanistic or structural insights are needed before we can fully understand how these proteins acted as the drivers of outer membrane protein assembly throughout this evolutionary scenario.

Ancestral Machines Have Been Modified and Recombined in the Course of Evolution

As the number of genes relocated from the endosymbiont to the host cell nucleus increased from tens to hundreds (and up >1,000), the need to translocate this vast repertoire of proteins would have substantially changed the demands placed on these former bacterial membranes. Novel translocation machinery therefore demanded innovation beyond the means of modified bacterial machinery. One way that mitochondria responded to this demand was to tinker with existing machinery, deploying it in new ways.

In many eukaryotes, there are two inner membrane translocases in mitochondria: TIM22 and TIM23 (Fig. 1). These translocases have divergent functions, with TIM22 responsible for delivering the many poly-topic membrane proteins of the inner membrane, and TIM23 handling the bulk of the matrix destined proteins as well as select inner membrane proteins. The channel components of TIM23 and TIM22 are related to each other by sequence and were derived from a gene duplication event (Cavalier-Smith 2006; Gross and Bhattacharya 2009; Jensen and Dunn 2002; Schneider et al. 2008). It can therefore be readily argued that the first mitochondria had a single TIM complex. At least two diverse sets of parasites, the Trypanosomes and the Microsporidia, contain only one TIM complex. While it is unclear at this point if Trypanosomes ancestrally possessed two TIMs, Microsporidia are related to Fungi providing clear evidence that in this case one of the duplicated TIMs has been secondarily lost. These parasites serve as an excellent example of how in an ancestral mitochondrion such a single TIM complex could provide both translocation and membrane assembly functions.

The core of the TIM23 complex is the Tim23 subunit, a poly-topic membrane protein that forms the protein import channel (Alder et al. 2008; Truscott et al. 2001). In a groundbreaking paper, Rassow et al. (1999) suggested that the Tim23 channel was derived from an amino acid transporter called LivH. While this ancestral LivH relationship has been questioned (Gross and Bhattacharya 2009), the model proposed by Rassow and colleagues remains enlightening: that a channel capable of transporting bulky, hydrophobic amino acids across the inner membrane would require relatively little adaptation in order to transfer polypeptides. The crux would be providing sufficient energy to “pull” a polymer of these amino acids through the channel. Even though the pair-wise sequence conservation between LivH family members and Tim23 proteins is very low (Cavalier-Smith 2006; Gross and Bhattacharya 2009; Murcha et al. 2007; Rassow et al. 1999), a signature PRAT (PReprotein and Amino acid Transporters) motif found in the Tim23 family of mitochondrial translocases is also present in the LivH protein of bacteria (Cavalier-Smith 2006; Murcha et al. 2007; Rassow et al. 1999). Systematic analysis of the family in *Arabidopsis thaliana* has also shown that plants retain mitochondrial versions of this protein family that transport amino acids as well as proteins (Murcha et al. 2007).

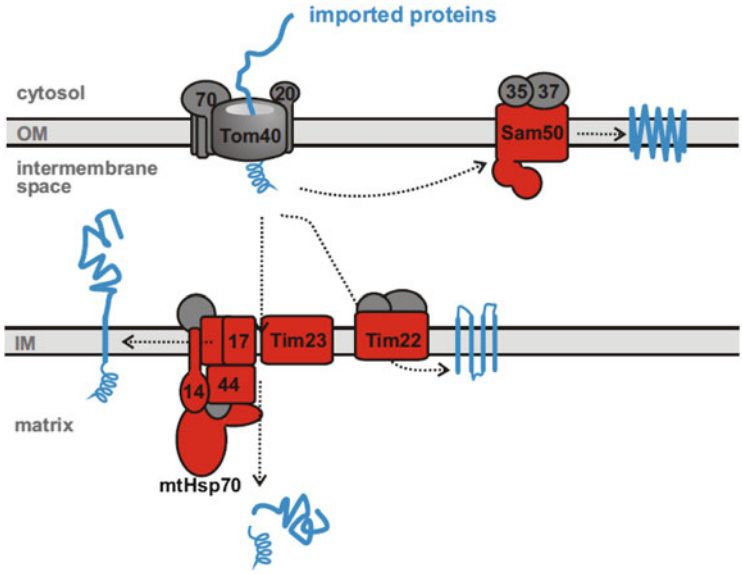


Fig. 2 Components of the major protein translocases in the mitochondrial inner and outer membranes. The details of the components of these machines have been recently reviewed (Chacinska et al. 2009; Neupert and Herrmann 2007). The major, essential component of each translocase is labelled (e.g. Tom40, Sam50, Tim23 and Tim22) and other subunits are given only numbers (e.g. "20" and "70" in the TOM complex refers to Tom20 and Tom70). For consistency, the *red* and *grey shading* denotes proteins that have or do not have identifiable bacterial ancestors, respectively. The functions of the various components are discussed in the text

The work required to drive vectorial polypeptide transport is provided by the import motor mtHsp70 (Fig. 2). This mtHsp70 is derived from a bacterial Hsp70 called DnaK, a protein found ubiquitously in bacteria (Boorstein et al. 1994). The import motor is docked to the TIM23 translocase by the Tim44 subunit (Rassow et al. 1994; Schneider et al. 1994) and a protein called Pam18/Tim14 regulates motor ATPase activity (D'Silva et al. 2003; Mokranjac et al. 2003; Truscott et al. 2003). Recent work has shown that α-proteobacteria possess inner membrane proteins with strong sequence similarity to the Tim44 (TimA) and Tim14 (TimB) core subunits of the TIM23 complex (Clements et al. 2009; Dolezal et al. 2006; Kutik et al. 2009) and a single point mutation in the *Caulobacter crescentus* Tim14 homologue is sufficient to convert it to a functional yeast TIM23 translocase subunit (Clements et al. 2009). These components were all available to evolution and, with even a rudimentary TIM complex in place [in the continued presence of both SecYEG and YidC translocases (Dolezal et al. 2006)], the proto-mitochondrion would have had a functional system for import of both matrix and inner membrane proteins. A primitive system such as this would provide the basis for the evolution of the highly specialised TIM translocases in extant organisms.

New Machines—Without a (Bacterial) Trace

The translocase of the outer membrane, or TOM complex, provides the general pore through which all proteins for the inner compartments of the mitochondrion (and also some outer membrane proteins) must pass (Fig. 1). Its development was key to the original bacterial endosymbiont transitioning into a genuine organelle. It has been suggested that the first protein translocase system in the "proto-mitochondrion" would have involved a primitive set-up: a β-barrel protein in the outer membrane and substrates in the host cytosol predisposed for targeting to mitochondria (Clements et al. 2009; Lucattini et al. 2004). Could such a simple pore have provided for the early needs of this primitive organelle? Microsporidians provide a true proof-of-principle example of such a simple TOM complex. These relatives of fungi possess readily identified TOM and TIM homologues (Burri et al. 2006; Waller et al. 2009). Only two TOM proteins are encoded in the complete genome of microsporidian *Encephalitozoon cuniculi*: Tom70 and Tom40. Given the function of Tom70 as a receptor, acting prior to the translocation reaction, this says that Tom40 alone can form a functional protein translocation pore. While this reduced TOM is clearly a result of secondary gene loss in microsporidians, it demonstrates the feasibility of a simpler primitive TOM complex in the ancestral endosymbiont.

Current phylogenetic analysis does not establish the ancestry of Tom40. Based on its predicted β-barrel topology it is broadly accepted that Tom40 was derived from the genome of the endosymbiont (Alcock et al. 2010; Bains and Lithgow 1999; Cavalier-Smith 2006; Gabriel et al. 2001; Herrmann 2003; Kutik et al. 2009; Mannella et al. 1996; Neupert and Herrmann 2007). Like all bacteria with two membranes, the endosymbiont would have had a range of β-barrel outer membrane proteins. Initially synthesised within the endosymbiont, a primitive TOM translocase could have been transported to the periplasm using the bacterial export pathway. If Tom40 was derived from a bacterial protein export channel it now imports, rather than exports, proteins but even this difference need not be problematic: biochemical analysis of purified mitochondrial outer membrane vesicles has shown that purified proteins can move in either direction through the TOM channel (Mayer et al. 1995).

The small TIM chaperones are found only in eukaryotes (Gentle et al. 2007), where they transfer precursors of both inner and outer mitochondrial membrane proteins from the TOM complex to the appropriate downstream machinery (either TIM22 or SAM, Fig. 1) (Hoppins and Nargang 2004; Jarosch et al. 1996; Koehler et al. 1998a, b; Wiedemann et al. 2004). Comparative analysis of the analogous but non-homologous prokaryotic chaperone (SurA) with the small TIMs (e.g. Tims 9 and 10) shows that while both chaperones can bind similar substrates, only the small TIM chaperones can transfer mitochondrial inner membrane proteins to the TIM22 translocase for insertion (Alcock et al. 2008). The small TIM family may therefore have arisen to enhance transport of inner membrane proteins, and also proved competent in transfer of outer membrane precursors,

leaving bacterial chaperones like SurA redundant. But how does a chaperone like the small TIM evolve, where apparently four distinct protein subunits are required? It starts from one. Again, an investigation of a parasite and its secondary reduction in genes provides the proof of principle for how a single gene product could give rise to the small TIM systems seen today. *C. parvum* is a single-celled human parasite possessing simple mitochondria with and a single small TIM protein that forms a homo-hexameric chaperone (Alcock et al. 2012).

As mentioned earlier in this review, the SAM complex can engage with two outer membrane proteins found only in fungi: Mdm10 and Mim1 (Thornton et al. 2010). Mdm10 is a modular component of two complexes which seem to function in distinct pathways for assembly of outer membrane proteins. A SAM-Mdm10 complex assists in assembly of the TOM complex. Mim1 is another modular subunit that can engage with the SAM complex (Becker et al. 2008). Mim1 functions in assembly of integral membrane proteins (Becker et al. 2011; Hulett et al. 2007; Ishikawa et al. 2004; Lueder and Lithgow 2009; Meisinger et al. 2004; Thornton et al. 2010). Mim1 and Mdm10 are each required for integration of different subunits into the TOM complex, demonstrating substrate specificity for each module, and Mim1 at least is involved in assembling other membrane proteins, including poly-topic proteins, into the outer membrane (Becker et al. 2011). Despite their functionally important and fundamental roles in assembling membrane proteins, neither Mdm10 nor Mim1 appears to be conserved outside the fungal lineage. This highlights the need to characterise protein import in organisms other than yeast: something analogous would likely function in place of Mdm10 and Mim1 in other organisms.

New Insights into Organelle Evolution from Mitochondrial Diversity

While clear insights into the adaptation of protein import pathways in the nascent mitochondrial organelle have been gained by studies of equivalent bacterial systems and minimalist mitochondria in parasites, to truly understand the evolution of mitochondrial import systems we must graft these insights onto the full diversity of the tree of eukaryotic life and test their universality. This is already providing new insights into, and questions regarding, mitochondrial import evolution. Figure 3 summarises the current state of knowledge of the presence of major components of mitochondrial import systems across the diversity of eukaryotes (Table 1 indicates module function with accessory/regulatory components shaded in grey). This snapshot considers detectible presence, rather than verifiable absence. The presence of a common core of import-related proteins found throughout eukaryotes points to the earliest innovations of this system after the establishment of the mitochondrion as a true organelle but before the radiation of present lineages

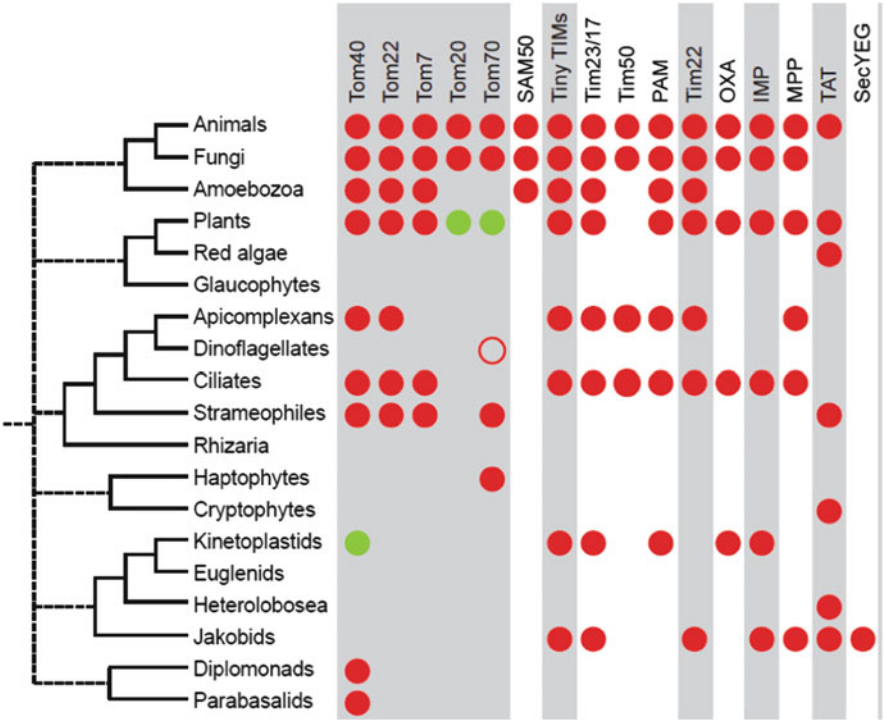


Fig. 3 Phylogeny of eukaryotes and known presences of mitochondrial import proteins and machines. *Red circles* indicate presence of homologues, *green circles* indicate presence of analogous proteins and *open circle* indicates indirect evidence based on function. *Dashed lines* indicate uncertainty in the eukaryotic phylogeny. Note: absences are generally not verified, and often represent missing data [references: Dagley et al. (2009), Danne and Waller (2011), Likic et al. (2010), Lister et al. (2003), Lithgow and Schneider (2010), Macasev et al. (2004), Perry et al. (2008), Pusnik et al. (2009), Schneider et al. (2008), Smith et al. (2007), Tong et al. (2011), Tsaousis et al. (2011), van Dooren et al. (2006), Wang and Lavrov (2007) and Yen et al. (2002)]

(Dolezal et al. 2006). Some newer insights challenging aspects of our views of mitochondrial import diversity and conservation are discussed in this section.

The TOM component Tom70 has been considered a specific receptor of the animal/fungal lineage where it handles inner membrane proteins with internal targeting sequences (Chan et al. 2006). The notion of Tom70 as a more recent mitochondrial innovation has been previously bolstered by failure to find Tom70 in the Amoebozoa (sister to animals/fungi), and the presence of an analogous receptor in plants, mtOM64, that is apparently derived from a chloroplast translocase in the absence of Tom70 (Chew et al. 2004; Perry et al. 2008). Unexpectedly, however, Tom70 homologues have recently been identified within distantly related groups to animals and fungi, the Stramenopiles and Haptophytes, including verification of heterologous function of one homologue in yeast (Tsaousis et al. 2011). Furthermore, studies of targeting signals in dinoflagellates indicate the functional

Table 1 Major known components of mitochondrial protein sorting machinery

	Modules	**Subunits in module** (yeast)	**Function of module**
TOM complex	Core translocase	Tom40, Tom22, Tom7	Translocation channel
	small subunits	Tom6, Tom5	Assists substrate transfer
	receptors	Tom70, Tom20	Promote substrate binding
SAM complex	Core translocase	Sam50	Membrane protein assembly
	metaxins	Sam35, Sam37	Assist protein assembly?
	Mdm10	Mdm10 (others?)	Assists protein assembly?
Tiny TIMs	Core complexes	Tim9, Tim10 and Tim8, Tim13	Transfer of substrates to TIM22 or SAM complexes
TIM22 complex	Core translocase	Tim22	Assembly of proteins into inner membrane.
	peripheral Tim accessory subunits	Tim12 Tim54, Tim18	Docking of tiny TIMs Assists protein assembly?
TIM23 complex	Core translocase	Tim23, Tim17	Translocation channel
	Tim50 PAM complex	Tim50 Pam18, Pam16, Tim44, mHsp70	Regulates channel opening Transfer of substrates into the matrix
	Tim21	Tim21	Regulates module docking
OXA complex	Core chaperone	Oxa1	Assembly of proteins into inner membrane.
	ribosome receptors	Mba1, Mdm38, Ylh47	Docking of mitochondrial ribosomes
IMP complex	Core peptidase	Imp1, Imp2	Processing of transfer-type sequences
	substrate binding	Som1	Modulates recognition
MPP	Core peptidase	Mas1, Mas2	Processing of N-terminal presequences in matrix
TAT	Core translocase	TatA	Translocation channel
	substrate binding	TatC	Promotes substrate binding
SecYEG	Core translocase	SecY	Translocation channel

conservation of the internal targeting signals between fungi and dinoflagellates, which further suggests that Tom70 function occurs very broadly, even though homologues are yet to be identified (Danne and Waller 2011). These observations suggest that Tom70 developed very early in mitochondrial evolution and has potentially been lost in several eukaryotic groups.

Two further examples of loss of early acquired targeting pathways reinforce that complexity gained during organelle evolution can also be reversed. As discussed above, there is compelling evidence of only a single TIM23/22 in both trypanosomatids and microsporidia (Schneider et al. 2008; Waller et al. 2009). It is unlikely that either of these two parasite groups have simply eliminated the substrates for one of these major complexes as both continue to import both polytopic membrane proteins and matrix proteins. Thus either TIM23 or TIM22 has apparently broadened its substrate range more recently in these lineages such that at least one of these inner membrane translocases became redundant and could be lost.

The bacterial SecYEG secretion route, on the other hand, is an ancestral pathway that likely played an early role in mitochondrial evolution but was then generally lost. Curiously it has been retained only in one group, the Jakobids (Lang et al. 1997). Why this group has continued to employ the SecY pathway, presumably for insertion of one or more proteins into the intermembrane space from the matrix, is unclear. Jakobids possess the most gene rich mitochondrial genomes of any mitochondria, and clues to SecYEG retention may be indicated by the presence of some of these genes. The *cox11* gene product, for example, has been suggested as a possible substrate for the SecY complex (Tong et al. 2011); however, this gene is present on at least one other mitochondrial genome, i.e. of the Heterolobosean, *Naegleria* (Gray et al. 2004). Thus, mitochondrial location of this gene alone does not necessarily require the maintenance of SecYEG.

A further observation that can be made from the diversity of mitochondrial targeting systems, and one that the models of fungi and animals alone have not illuminated, is the broad retention of the twin-arginine translocation (Tat) pathway. The Tat pathway is a bacterial secretion system that can translocate folded proteins as well as proteins complexed with cofactors or even other proteins (Berks et al. 2005; Natale et al. 2008). In bacteria, Tat substrates include many redox-related proteins such as respiratory proteins that contain cofactors necessary for electron transport processes. The Tat pathway has also been inherited by plastids where it is involved in protein transport into the internal thylakoid membranes (and presumed not to function in the organelle envelope membranes) (Müller and Klösgen 2005). A Tat pathway has not been investigated in mitochondria, although the presence of genes for Tat translocator components has often been noted in mitochondrial genomes. Figure 3 indicates major eukaryotic groups where Tat genes (also known as *ymf16* and *mttB*) are found, and this includes basal animal lineages (choanoflagellates and some sponges), plants and several protist groups (Bogsch et al. 1998; Burger et al. 2003; Gray et al. 1999; Lang et al. 1997; Wang and Lavrov 2007, 2008; Yen et al. 2002). TatC is the most common gene indicator of this pathway in mitochondrial genomes, but Jakobids possess both TatC and TatA that potentially encode a minimal functional translocon. Presumably, TatA genes have relocated to the nucleus of most groups, and TatC has also been found in the nuclei of some basal animals (Wang and Lavrov 2007), indicating that nucleus-encoded genes might reveal an even wider occurrence of mitochondrial Tat.

The questions of what substrates might use a Tat pathway in mitochondria, and why or how (or even if) some groups have eliminated this pathway, are now in need of attention. There is clear scope for several cofactor containing redox-related proteins to use this pathway in mitochondria, potentially including substrates directly inherited from the bacterial progenitor of this organelle. These substrates might be proteins either still coded for on mitochondrial genomes, or for which the genes now occur in the nucleus as is also the case for many plastidal Tat substrates (Müller and Klösgen 2005). A possible way to identify them is to search mitochondrial proteins for the Tat pathway sorting signal, which resembles the Sec-type cleavable N-terminal signal sequence but also including a double arginine residue typically in the context Z-R-R-x-Φ-Φ (where Z stands for any polar residue and Φ

for hydrophobic residues) (Berks et al. 2005; Natale et al. 2008). While these bacterial signals are generally well conserved in the plastid system, some substrates are known to have more cryptic signals, and it is possible that a mitochondrial system has also diverged (Müller and Klösgen 2005). In plastids one known Tat substrate is the Fe–S containing Rieske protein of the cytochrome b_6f-complex (Rip1), and this protein also occurs in the equivalent mitochondrial cytochrome bc_1 complex (Molik et al. 2001). Intriguingly, in yeast this mitochondrial protein contains two N-terminal pre-peptides, one that acts as a mitochondrial matrix-targeting peptide, and the second one that is removed upon final delivery back across the inner membrane to the intermembrane space (Conte and Zara 2011). The significance of this second peptide, and the route of this final targeting event, is unknown.[1] Such a protein might offer scope to explore the role of the mitochondrial Tat pathway in either this or another experimentally tractable system.

The broad themes of re-development and re-deployment discussed in sections "An Ancestral System Doing (More or Less) What It Has Always Done" and "Ancestral Machines Have Been Modified and Recombined in the Course of Evolution" showcase the first chapters in our understanding of the evolution of the mitochondrial import machinery (Hewitt et al. 2011). Where details of the transition from endosymbiont to organelle have been obscured, section "New Machines—Without a (Bacterial) Trace" shows how minimalist mitochondrial systems of parasites provide a valuable testing ground for hypotheses regarding the earliest stages of mitochondrial evolution. Ultimately our understanding of organelle evolution should encompass the whole eukaryotic tree, and with the advent of next-gen sequencing methodologies, genome sequencing is economically and practically achievable for even unculturable and otherwise obscure organisms. The re-evaluation of Tom70 in light of newly identified homologues is an example of how bioinformatics alone can reshape our understanding of the evolutionary history of protein transport machinery. However, as with most of the insights discussed here, the full power of these bioinformatics discoveries is realised through biochemical verification and further investigation of the questions raised by an expanding view of mitochondrial diversity.

References

Alcock FH, Grossmann JG, Gentle IE, Likic VA, Lithgow T, Tokatlidis K (2008) Conserved substrate binding by chaperones in the bacterial periplasm and the mitochondrial intermembrane space. Biochem J 409:377–387

[1] Note added in proof: The assembly of Rip1 was recently investigated and found to depend on the AAA-ATPase Bcs1 (Wagener et al. 2011). This raises the intriguing possibility that the Tat pathway has been replaced by a functionally homologous system that translocates the folded Fe–S containing domain, but uses an entirely different protein3 machinery to do so.

Alcock F, Clements A, Webb C, Lithgow T (2010) Tinkering inside the organelle. Science 327:649–650

Alcock F, Webb CT, Dolezal P, Hewitt V, Shingu-Vasquez M, Likic VA, Traven A, Lithgow T (2012) A small Tim homohexamer in the relict mitochondrion of *Cryptosporidium*. Mol Biol Evol 29:113–122

Alder NN, Jensen RE, Johnson AE (2008) Fluorescence mapping of mitochondrial TIM23 complex reveals a water-facing, substrate-interacting helix surface. Cell 134:439–450

Andersson SG, Kurland CG (1999) Origins of mitochondria and hydrogenosomes. Curr Opin Microbiol 2:535–541

Bains G, Lithgow T (1999) The Tom channel in the mitochondrial outer membrane: alive and kicking. Bioessays 21:1–4

Becker T, Pfannschmidt S, Guiard B, Stojanovski D, Milenkovic D, Kutik S, Pfanner N, Meisinger C, Wiedemann N (2008) Biogenesis of the mitochondrial TOM complex: Mim1 promotes insertion and assembly of signal-anchored receptors. J Biol Chem 283:120–127

Becker T, Wenz L-S, Krüger V, Lehmann W, Müller JM, Goroncy L, Zufall N, Lithgow T, Guiard B, Chacinska A, Wagner R, Meisinger C, Pfanner N (2011) The mitochondrial import protein Mim1 promotes biogenesis of multispanning outer membrane proteins. J Cell Biol 194:387–395

Berks BC, Palmer T, Sargent F (2005) Protein targeting by the bacterial twin-arginine translocation (Tat) pathway. Curr Opin Microbiol 8:174–181

Bogsch EG, Sargent F, Stanley NR, Berks BC, Robinson C, Palmer T (1998) An essential component of a novel bacterial protein export system with homologues in plastids and mitochondria. J Biol Chem 273:18003–18006

Boorstein WR, Ziegelhoffer T, Craig EA (1994) Molecular evolution of the HSP70 multigene family. J Mol Evol 38:1–17

Burger G, Forget L, Zhu Y, Gray MW, Lang BF (2003) Unique mitochondrial genome architecture in unicellular relatives of animals. Proc Natl Acad Sci USA 100:892–897

Burri L, Williams BA, Bursac D, Lithgow T, Keeling PJ (2006) Microsporidian mitosomes retain elements of the general mitochondrial targeting system. Proc Natl Acad Sci USA 103:15916–15920

Cavalier-Smith T (2002) The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa. Int J Syst Evol Microbiol 52:297–354

Cavalier-Smith T (2006) Rooting the tree of life by transition analyses. Biol Direct 1:19

Chacinska A, Koehler CM, Milenkovic D, Lithgow T, Pfanner N (2009) Importing mitochondrial proteins: machineries and mechanisms. Cell 138:628–644

Chan NC, Likic VA, Waller RF, Mulhern TD, Lithgow T (2006) The C-terminal TPR domain of Tom70 defines a family of mitochondrial protein import receptors found only in animals and fungi. J Mol Biol 358:1010–1022

Chew O, Lister R, Qbadou S, Heazlewood JL, Soll J, Schleiff E, Millar AH, Whelan J (2004) A plant outer mitochondrial membrane protein with high amino acid sequence identity to a chloroplast protein import receptor. FEBS Lett 557:109–114

Clements A, Bursac D, Gatsos X, Perry AJ, Civciristov S, Celik N, Likic VA, Poggio S, Jacobs-Wagner C, Strugnell RA, Lithgow T (2009) The reducible complexity of a mitochondrial molecular machine. Proc Natl Acad Sci USA 106:15791–15795

Conte L, Zara V (2011) The rieske iron-sulfur protein: import and assembly into the cytochrome bc(1) complex of yeast mitochondria. Bioinorg Chem Appl 2011:363941

D'Silva PD, Schilke B, Walter W, Andrew A, Craig EA (2003) J protein cochaperone of the mitochondrial inner membrane required for protein import into the mitochondrial matrix. Proc Natl Acad Sci USA 100:13839–13844

Dagley MJ, Dolezal P, Likic VA, Smid O, Purcell AW, Buchanan SK, Tachezy J, Lithgow T (2009) The protein import channel in the outer mitosomal membrane of *Giardia intestinalis*. Mol Biol Evol 26:1941–1947

Danne JC, Waller RF (2011) Analysis of dinoflagellate mitochondrial protein sorting signals indicates a highly stable protein targeting system across eukaryotic diversity. J Mol Biol 408:643–653

de Duve C (2007) The origin of eukaryotes: a reappraisal. Nat Rev Genet 8:395–403

Dolezal P, Likic V, Tachezy J, Lithgow T (2006) Evolution of the molecular machines for protein import into mitochondria. Science 313:314–318

Driessen AJ, Manting EH, van der Does C (2001) The structural basis of protein targeting and translocation in bacteria. Nat Struct Biol 8:492–498

Embley TM, Martin W (2006) Eukaryotic evolution, changes and challenges. Nature 440:623–630

Funes S, Hasona A, Bauerschmitt H, Grubbauer C, Kauff F, Collins R, Crowley PJ, Palmer SR, Brady LJ, Herrmann JM (2009) Independent gene duplications of the YidC/Oxa/Alb3 family enabled a specialized cotranslational function. Proc Natl Acad Sci USA 106:6656–6661

Gabaldon T, Huynen MA (2007) From endosymbiont to host-controlled organelle: the hijacking of mitochondrial protein synthesis and metabolism. PLoS Comput Biol 3:e219

Gabriel K, Buchanan SK, Lithgow T (2001) The alpha and the beta: protein translocation across mitochondrial and plastid outer membranes. Trends Biochem Sci 26:36–40

Gatsos X, Perry AJ, Anwari K, Dolezal P, Wolynec PP, Likic VA, Purcell AW, Buchanan SK, Lithgow T (2008) Protein secretion and outer membrane assembly in Alphaproteobacteria. FEMS Microbiol Rev 32:995–1009

Gentle I, Gabriel K, Beech P, Waller R, Lithgow T (2004) The Omp85 family of proteins is essential for outer membrane biogenesis in mitochondria and bacteria. J Cell Biol 164:19–24

Gentle IE, Burri L, Lithgow T (2005) Molecular architecture and function of the Omp85 family of proteins. Mol Microbiol 58:1216–1225

Gentle IE, Perry AJ, Alcock FH, Likic VA, Dolezal P, Ng ET, Purcell AW, McConnville M, Naderer T, Chanez AL, Charriere F, Aschinger C, Schneider A, Tokatlidis K, Lithgow T (2007) Conserved motifs reveal details of ancestry and structure in the small TIM chaperones of the mitochondrial intermembrane space. Mol Biol Evol 24:1149–1160

Gray MW, Burger G, Lang BF (1999) Mitochondrial evolution. Science 283:1476–1481

Gray MW, Burger G, Lang BF (2001) The origin and early evolution of mitochondria. Genome Biol 2:Reviews1018

Gray MW, Lang BF, Burger G (2004) Mitochondria of protists. Annu Rev Genet 38:477–524

Gross J, Bhattacharya D (2009) Mitochondrial and plastid evolution in eukaryotes: an outsiders' perspective. Nat Rev Genet 10:495–505

Herrmann JM (2003) Converting bacteria to organelles: evolution of mitochondrial protein sorting. Trends Microbiol 11:74–79

Hewitt V, Alcock F, Lithgow T (2011) Minor modifications and major adaptations: the evolution of molecular machines driving mitochondrial protein import. Biochim Biophys Acta 1808:947–954

Hoppins SC, Nargang FE (2004) The Tim8-Tim13 complex of *Neurospora crassa* functions in the assembly of proteins into both mitochondrial membranes. J Biol Chem 279:12396–12405

Hulett JM, Walsh P, Lithgow T (2007) Domain stealing by receptors in a protein transport complex. Mol Biol Evol 24:1909–1911

Ishikawa D, Yamamoto H, Tamura Y, Moritoh K, Endo T (2004) Two novel proteins in the mitochondrial outer membrane mediate beta-barrel protein assembly. J Cell Biol 166:621–627

Jarosch E, Tuller G, Daum G, Waldherr M, Voskova A, Schweyen RJ (1996) Mrs5p, an essential protein of the mitochondrial intermembrane space, affects protein import into yeast mitochondria. J Biol Chem 271:17219–17225

Jensen RE, Dunn CD (2002) Protein import into and across the mitochondrial inner membrane: role of the TIM23 and TIM22 translocons. Biochim Biophys Acta 1592:25–34

Kiefer D, Kuhn A (2007) YidC as an essential and multifunctional component in membrane protein assembly. Int Rev Cytol 259:113–138

Knowles TJ, Scott-Tucker A, Overduin M, Henderson IR (2009) Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. Nat Rev Microbiol 7:206–214

Koehler CM, Jarosch E, Tokatlidis K, Schmid K, Schweyen RJ, Schatz G (1998a) Import of mitochondrial carriers mediated by essential proteins of the intermembrane space. Science 279:369–373

Koehler CM, Merchant S, Oppliger W, Schmid K, Jarosch E, Dolfini L, Junne T, Schatz G, Tokatlidis K (1998b) Tim9p, an essential partner subunit of Tim10p for the import of mitochondrial carrier proteins. EMBO J 17:6477–6486

Kutik S, Stroud DA, Wiedemann N, Pfanner N (2009) Evolution of mitochondrial protein biogenesis. Biochim Biophys Acta 1790:409–415

Lang BF, Burger G, O'Kelly CJ, Cedergren R, Golding GB, Lemieux C, Sankoff D, Turmel M, Gray MW (1997) An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. Nature 387:493–497

Lang BF, Gray MW, Burger G (1999) Mitochondrial genome evolution and the origin of eukaryotes. Annu Rev Genet 33:351–397

Likic VA, Dolezal P, Celik N, Dagley M, Lithgow T (2010) Using hidden Markov models to discover new protein transport machines. Methods Mol Biol 619:271–284

Lister R, Murcha MW, Whelan J (2003) The mitochondrial protein import machinery of plants (MPIMP) database. Nucleic Acids Res 31:325–327

Lithgow T, Schneider A (2010) Evolution of macromolecular import pathways in mitochondria, hydrogenosomes and mitosomes. Philos Trans R Soc Lond B Biol Sci 365:799–817

Lucattini R, Likic VA, Lithgow T (2004) Bacterial proteins predisposed for targeting to mitochondria. Mol Biol Evol 21:652–658

Lueder F, Lithgow T (2009) The three domains of the mitochondrial outer membrane protein Mim1 have discrete functions in assembly of the TOM complex. FEBS Lett 583:1475–1480

Macasev D, Whelan J, Newbigin E, Silva-Filho MC, Mulhern TD, Lithgow T (2004) Tom22', an 8-kDa trans-site receptor in plants and protozoans, is a conserved feature of the TOM complex that appeared early in the evolution of eukaryotes. Mol Biol Evol 21:1557–1564

Mannella CA, Neuwald AF, Lawrence CE (1996) Detection of likely transmembrane beta strand regions in sequences of mitochondrial pore proteins using the Gibbs sampler. J Bioenerg Biomembr 28:163–169

Mayer A, Neupert W, Lill R (1995) Mitochondrial protein import: reversible binding of the presequence at the trans side of the outer membrane drives partial translocation and unfolding. Cell 80:127–137

Meisinger C, Rissler M, Chacinska A, Szklarz LK, Milenkovic D, Kozjak V, Schonfisch B, Lohaus C, Meyer HE, Yaffe MP, Guiard B, Wiedemann N, Pfanner N (2004) The mitochondrial morphology protein Mdm10 functions in assembly of the preprotein translocase of the outer membrane. Dev Cell 7:61–71

Mokranjac D, Sichting M, Neupert W, Hell K (2003) Tim14, a novel key component of the import motor of the TIM23 protein translocase of mitochondria. EMBO J 22:4945–4956

Molik S, Karnauchov I, Weidlich C, Herrmann RG, Klosgen RB (2001) The Rieske Fe/S protein of the cytochrome b6/f complex in chloroplasts: missing link in the evolution of protein transport pathways in chloroplasts? J Biol Chem 276:42761–42766

Müller M, Klösgen RB (2005) The Tat pathway in bacteria and chloroplasts (review). Mol Membr Biol 22:113–121

Murcha MW, Elhafez D, Lister R, Tonti-Filippini J, Baumgartner M, Philippar K, Carrie C, Mokranjac D, Soll J, Whelan J (2007) Characterization of the preprotein and amino acid transporter gene family in *Arabidopsis*. Plant Physiol 143:199–212

Natale P, Bruser T, Driessen AJ (2008) Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane – distinct translocases and mechanisms. Biochim Biophys Acta 1778:1735–1756

Neupert W, Herrmann JM (2007) Translocation of proteins into mitochondria. Annu Rev Biochem 76:723–749

Ott M, Herrmann JM (2010) Co-translational membrane insertion of mitochondrially encoded proteins. Biochim Biophys Acta 1803:767–775

Perry AJ, Rimmer KA, Mertens HD, Waller RF, Mulhern TD, Lithgow T, Gooley PR (2008) Structure, topology and function of the translocase of the outer membrane of mitochondria. Plant Physiol Biochem 46:265–274

Pohlschroder M, Gimenez MI, Jarrell KF (2005) Protein transport in Archaea: Sec and twin arginine translocation pathways. Curr Opin Microbiol 8:713–719

Pusnik M, Charriere F, Maser P, Waller RF, Dagley MJ, Lithgow T, Schneider A (2009) The single mitochondrial porin of *Trypanosoma brucei* is the main metabolite transporter in the outer mitochondrial membrane. Mol Biol Evol 26:671–680

Rassow J, Maarse AC, Krainer E, Kubrich M, Muller H, Meijer M, Craig EA, Pfanner N (1994) Mitochondrial protein import: biochemical and genetic evidence for interaction of matrix hsp70 and the inner membrane protein MIM44. J Cell Biol 127:1547–1556

Rassow J, Dekker PJ, van Wilpe S, Meijer M, Soll J (1999) The preprotein translocase of the mitochondrial inner membrane: function and evolution. J Mol Biol 286:105–120

Ruiz N, Kahne D, Silhavy TJ (2006) Advances in understanding bacterial outer-membrane biogenesis. Nat Rev Microbiol 4:57–66

Samuelson JC, Chen M, Jiang F, Moller I, Wiedmann M, Kuhn A, Phillips GJ, Dalbey RE (2000) YidC mediates membrane protein insertion in bacteria. Nature 406:637–641

Schneider HC, Berthold J, Bauer MF, Dietmeier K, Guiard B, Brunner M, Neupert W (1994) Mitochondrial Hsp70/MIM44 complex facilitates protein import. Nature 371:768–774

Schneider A, Bursac D, Lithgow T (2008) The direct route: a simplified pathway for protein import into the mitochondrion of trypanosomes. Trends Cell Biol 18:12–18

Smith DG, Gawryluk RM, Spencer DF, Pearlman RE, Siu KW, Gray MW (2007) Exploring the mitochondrial proteome of the ciliate protozoon *Tetrahymena thermophila*: direct analysis by tandem mass spectrometry. J Mol Biol 374:837–863

Stuart R (2002) Insertion of proteins into the inner membrane of mitochondria: the role of the Oxa1 complex. Biochim Biophys Acta 1592:79–87

Thornton N, Stroud DA, Milenkovic D, Guiard B, Pfanner N, Becker T (2010) Two modular forms of the mitochondrial sorting and assembly machinery are involved in biogenesis of alpha-helical outer membrane proteins. J Mol Biol 396:540–549

Timmis JN, Ayliffe MA, Huang CY, Martin W (2004) Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nat Rev Genet 5:123–135

Tong J, Dolezal P, Selkrig J, Crawford S, Simpson AG, Noinaj N, Buchanan SK, Gabriel K, Lithgow T (2011) Ancestral and derived protein import pathways in the mitochondrion of *Reclinomonas americana*. Mol Biol Evol 28:1581–1591

Truscott KN, Kovermann P, Geissler A, Merlin A, Meijer M, Driessen AJ, Rassow J, Pfanner N, Wagner R (2001) A presequence- and voltage-sensitive channel of the mitochondrial preprotein translocase formed by Tim23. Nat Struct Biol 8:1074–1082

Truscott KN, Voos W, Frazier AE, Lind M, Li Y, Geissler A, Dudek J, Muller H, Sickmann A, Meyer HE, Meisinger C, Guiard B, Rehling P, Pfanner N (2003) A J-protein is an essential subunit of the presequence translocase-associated protein import motor of mitochondria. J Cell Biol 163:707–713

Tsaousis AD, Gaston D, Stechmann A, Walker PB, Lithgow T, Roger AJ (2011) A functional Tom70 in the human parasite *Blastocystis* sp.: implications for the evolution of the mitochondrial import apparatus. Mol Biol Evol 28:781–791

van der Laan M, Nouwen NP, Driessen AJ (2005) YidC – an evolutionary conserved device for the assembly of energy-transducing membrane protein complexes. Curr Opin Microbiol 8:182–187

van Dooren GG, Stimmler LM, McFadden GI (2006) Metabolic maps and functions of the *Plasmodium* mitochondrion. FEMS Microbiol Rev 30:596–630

Wagener N, Ackermann M, Funes S, Neupert W (2011) A pathway of protein translocation in mitochondria mediated by the AAA-ATPase Bcs1. Mol Cell 44:191–202

Waller RF, Jabbour C, Chan NC, Celik N, Likic VA, Mulhern TD, Lithgow T (2009) Evidence of a reduced and modified mitochondrial protein import apparatus in microsporidian mitosomes. Eukaryot Cell 8:19–26

Wang X, Lavrov DV (2007) Mitochondrial genome of the homoscleromorph *Oscarella carmela* (Porifera, Demospongiae) reveals unexpected complexity in the common ancestor of sponges and other animals. Mol Biol Evol 24:363–373

Wang X, Lavrov DV (2008) Seventeen new complete mtDNA sequences reveal extensive mitochondrial genome evolution within the Demospongiae. PLoS One 3:e2723

Webb CT, Lithgow T (2010) Mitochondrial biogenesis: sorting mechanisms cooperate in ABC transporter assembly. Curr Biol 20:R564–R567

Wiedemann N, Truscott KN, Pfannschmidt S, Guiard B, Meisinger C, Pfanner N (2004) Biogenesis of the protein import channel Tom40 of the mitochondrial outer membrane: intermembrane space components are involved in an early stage of the assembly pathway. J Biol Chem 279:18188–18194

Yen MR, Tseng YH, Nguyen EH, Wu LF, Saier MH Jr (2002) Sequence and phylogenetic analyses of the twin-arginine targeting (Tat) protein export system. Arch Microbiol 177:441–450

Part II
Autotrophy as the Driving Force for Endosymbiosis: Primary Endosymbiosis

The Single Primary Endosymbiotic Event

Wolfgang Löffelhardt

Abstract Eukaryotic phototrophs arose between about 1,600 and 1,200 Mya through the incorporation of a cyanobacterium by a phagotrophic eukaryote. In a very special and complex process, the cyanobacterium and the heterotrophic cell complemented each other that well to change the predator–prey relationship to a mutualistic one: the cyanobacterium was converted into an obligate endosymbiont allowing phototrophy of the host cell and ultimately into an organelle, the plastid. Pros and cons of a scenario assuming a single primary endosymbiotic event leading to a protoalga ancestral to the kingdom "Plantae" are discussed.

Contents

Introduction 39
Endosymbiotic Gene Transfer 41
Insights from Completely Sequenced Plastid Genomes 43
Outcome of Phylogenetic Analyses 44
Insights from Comparing the Protein Import Machineries of Plastids 45
The Cyanobacterial Ancestor 47
Alternative Scenarios 48
Conclusions 48
..... 48

Introduction

The striking similarity of cyanobacteria (then called blue-green algae) and plant chloroplasts with respect to morphology, pigmentation, and photosynthetic performance led Konstantin Mereschkowsky more than 100 years ago to propose

W. Löffelhardt (✉)
Department of Biochemistry and Cell Biology, University of Vienna, Dr. Bohrgasse 9, 1030 Vienna, Austria
e-mail: wolfgang.loeffelhardt@univie.ac.at

W. Löffelhardt (ed.), *Endosymbiosis*, DOI 10.1007/978-3-7091-1303-5_3,

© Springer-Verlag Wien 2014

"symbiogenesis" as an evolutionary link between organism and organelle (Mereschkowsky 1905). Later, a similar hypothesis was also formulated for the origin of mitochondria from endosymbiotic aerobic bacteria (Wallin 1925). Both concepts met only limited success and were nearly forgotten (Hagemann 2007) when at the beginning of the 1960s, organelle DNA was found in chloroplasts (Ris and Plaut 1962) and then also in mitochondria. This prompted Lynn Margulis–Sagan to fight for a revival of the "serial endosymbiosis hypothesis" (Sagan 1967; Margulis 1981). It is her merit to have pursued this idea despite objections from opponents (supporting, e.g., the "direct filiation hypothesis") for almost 20 years until its general acceptance. The pronounced differences in plastid morphology and pigmentation led to an early proposal for a polyphyletic evolution of plastids from different bacterial endosymbionts (Raven 1970). With the advent of molecular biology, comparison of 16S rRNA and, later on, protein sequences convincingly supported the monophyly of plastids and their cyanobacterial ancestry [for reviews, see Gould et al. (2008) and Keeling (2010)]. Clearly, mitochondria appeared first. However, it is not yet completely settled (Lang 2013) if these originated from an endosymbiosis between an α-proteobacterium and an amitochondriate protoeukaryote or rather through merging of two prokaryotes: an α-proteobacterium and an archaebacterium (Thiergart et al. 2011; Williams et al. 2012; Martijn and Ettema 2013). In the plant lineage, the acquisition of phototrophy through uptake of a cyanobacterium by a heterotrophic protist is called the "primary endosymbiotic event" (Fig. 1). This distinguishes it from "secondary endosymbiotic events," where eukaryotic algae became the complex plastids of several protist lineages through secondary symbioses. For mitochondria no such distinction is needed, because no examples of secondary mitochondria are known.

After copious adaptations and rearrangements an obligate endosymbiont was generated that underwent massive genome reduction (>95 %). From there the lineages known as Archaeplastida (Adl et al. 2005) branched off: glaucophytes, rhodophytes, and chlorophytes/streptophytes. While the photosynthetic apparatus of all primary plastids is very similar to that of cyanobacteria, that of muroplasts and rhodoplasts is almost identical to it. Unique cyanobacterial features of muroplasts are represented by the peptidoglycan wall (gray layer in Fig. 1) and the lack of Lhca (as depicted by the blue coloring of the stroma in Fig. 1).

Did primary endosymbiosis happen only once, i.e., was there but a single event leading to successful completion of an enormously complex integration process between a distinguished cyanobacterium and a distinguished heterotrophic protist? The resulting "protoalga" (the most ancient phototrophic eukaryote) would then give rise to all other algae and plants. Alternatively, multiple events under participation of different cyanobacteria and different eukaryotic hosts can be envisaged, followed by convergent evolution. The intrinsic problem is that experimental evidence or proof in favor of or against a single primary endosymbiotic event is not possible per se (Howe et al. 2008). One can only weigh the facts available to date and then decide which scenario seems more likely. The majority of researchers in the field (including the author of this chapter) took sides for the single event.

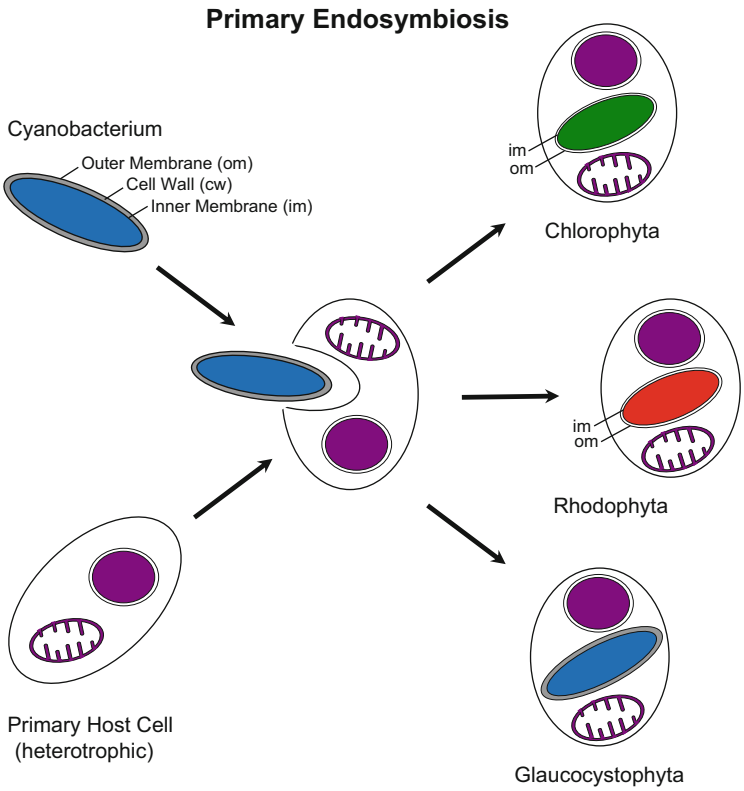


Fig. 1 Primary endosymbiosis. A cyanobacterium was engulfed by a protist through phagocytosis and underwent massive genome reduction. From this endosymbiont the plastids of glaucophytes, rhodophytes, and chlorophytes/streptophytes evolved, all surrounded by two membranes (Cavalier-Smith 2000). Modified after Stoebe and Maier (2002)

However, the reservations of some colleagues (Stiller 2007; Howe et al. 2008) should not be neglected. Despite those vagaries of gene tree comparison, the unity of the plastid protein import apparatus among all lineages with primary plastids remains strong evidence in favor of a single origin of plastids (see below).

Endosymbiotic Gene Transfer

The key process certainly is "endosymbiotic gene transfer" (EGT; Martin et al. 1993; Timmis et al. 2004) from the genome of the engulfed cyanobacterium to the nuclear genome of the protist host cell. EGT can be seen as a special form of (concerted) lateral (LGT) or horizontal gene transfer (HGT), different from vertical inheritance. Initially, the goal obviously was to prevent the endosymbiont from

escaping after he had proved to be useful. Considerable genome erosion of the cyanobacterial symbiont in the *Nostoc*/*Azolla* symbiosis was recently reported (Ran et al. 2010). During primary endosymbiosis, EGT served to gradually increase host nuclear dominance and control over the endosymbiont proteome. In a long-lasting process via semiautonomous, obligate endosymbionts the present organelle status was reached. The situation at the onset was very complicated: thousands of prokaryotic genes had to be transferred and had to find eukaryotic promoters and suitable 3′-UTRs allowing for their expression. In many cases, an upstream DNA sequence was needed functioning as a sorting signal, i.e., a stroma-targeting peptide (STP) directing the respective gene products back to the endosymbiont/organelle: compartmentation of photosynthesis and many other anabolic pathways did not change much compared to the cyanobacterial ancestor as illustrated by the quotation "plants are, biochemically, cyanobacteria wrapped in a bigger box" (Dagan et al. 2013). The reduction of the plastid genome to 2–5 % of that of the cyanobacterial ancestor largely occurred during the evolution of the protoalga and continues up to now in the plant kingdom, albeit at a very low rate. Complete loss of the plastid genome is not desirable: the expression of plastid-resident key photosynthesis genes must remain subjected to direct redox control through the electron transport chain in the same compartment (Allen et al. 2011). DNA transfer itself occurs after endosymbiont/organelle lysis, is frequent, and is, most likely, in "big chunks" (Huang et al. 2003; Martin 2003). The mechanism involved is nonhomologous recombination at double strand breaks in nuclear DNA (Leister and Kleine 2011). Consequently, nonfunctional nuclear plastid-like DNA sequences (nupts) are found in the nuclei of all higher plants investigated. In rice, chloroplast DNA is constantly taken up, shuffled, and degraded (Matsuo et al. 2005). In tobacco, there are also nupts with intact open reading frames waiting for several million years to become functional (Rousseau-Gueutin et al. 2011). In algae as, e.g., *Chlamydomonas reinhardtii* where there is but one chloroplast per cell, DNA transfer to the nucleus is much less frequent (Lister et al. 2003; Smith et al. 2011). An involvement of RNA/cDNA intermediates in plastid to nucleus gene transfer (as reported for some mitochondrial genes) could not be demonstrated despite comprehensive and well-designed genetic screens in tobacco with group II intron-containing chloroplast transgenes (Fuentes et al. 2012). In the laboratory, functional DNA transfer to the nucleus seems to work well for engineered chloroplasts with built-in selection (Stegemann and Bock 2006; Lloyd and Timmis 2011). However, EGT of let's say 2,500–3,000 "normal" genes in vivo with a return option for the gene products (accounting in number for a chloroplast proteome; Armbruster et al. 2011) can be expected a very lengthy process, as seen with the examples of more recent "primary" endosymbioses discussed below. The finding of a STP is considered an easier task than gene activation based on older data with mitochondrial import (Baker and Schatz 1987). This might not always be the case given that, e.g., Calvin cycle enzymes are abundant proteins where efficient import is critical. In such precursors, the STPs are kind of "canonical," i.e., better predictable than for average chloroplast proteins (Zybailov et al. 2008). In some proteomic studies, a number of soluble plastid proteins show up where the genes lack a

recognizable STP. Often, a signal peptide-like targeting signal is invoked in such cases. However, no evidence for a significant contribution of an ER-based (secretory) pathway for plastid import was found among 900 bona fide plastid proteins (Zybailov et al. 2008).

Insights from Completely Sequenced Plastid Genomes

The number and type of protein genes in the plastid genomes of phototrophic organisms are much more conserved than for mitochondria. The observed set of 46 genes common to all plastid types is indicative of a single event at the first glance but could also be the result of convergent evolution (Stiller et al. 2003). While plastome organization varies to some extent, especially between different algal phyla, several gene clusters known from cyanobacteria persist (with respect to gene order but with reduction in gene number due to EGT to the nucleus), e.g., the large ribosomal protein gene cluster (Stoebe and Kowallik 1999). By mid of the nineties plastid genomes of several green algae/higher plants, the red alga *Porphyra purpurea* and the glaucophyte *Cyanophora paradoxa* were completely sequenced, revealing—in addition to known cyanobacterial gene clusters/operons—conserved transcription units not present in cyanobacteria. One peculiar cluster contains genes for subunits of three different oligomeric enzymes or macromolecular complexes, i.e., RNA polymerase (PEP), ribosomes, and ATP synthase (Fig. 2). A cotranscription of all these genes with the purpose of producing stoichiometric amounts of the respective gene products is not required. Hence, there should be no selection pressure on the formation of such a cluster. On the contrary, in some chloroplasts, e.g., those of *Chlamydomonas reinhardtii* known for numerous rearrangements, this cluster is disrupted. These findings led to the denomination "gene cluster diagnostic for plastid evolution" and to the first formulations of a "single primary endosymbiotic event" (Reith and Munholland 1993; Kowallik 1994; Löffelhardt et al. 1997). At that time, just one cyanobacterial genome, that of *Synechocystis* sp. PCC6803, was sequenced. An early objection was that other cyanobacterial species, eventually some close to the plastid ancestor, might have a similar genome organization displaying this peculiar gene cluster. Now it is obvious that among the 34 complete cyanobacterial genomes listed in "cyanobase" (http://genome.kazusa.or.jp/cyanobase/), such a gene cluster is never found! Thus, the most likely explanation for the appearance of this very gene cluster in the genome of muroplasts, rhodoplasts, and chloroplasts is not convergent evolution but inheritance from the protoalga. At the stage of the protoplastid, the main part of endosymbiotic gene transfer to the nucleus was already performed resulting in a thorough shuffling of transcription units. Thereby, fortuitously, the diagnostic gene cluster was created. Most probably, it was very similar to the counterpart from extant rhodoplasts. In the following, individual, less pronounced gene losses occurred in the different lineages of algae (Fig. 2; Martin et al. 1998) and this process still continues.

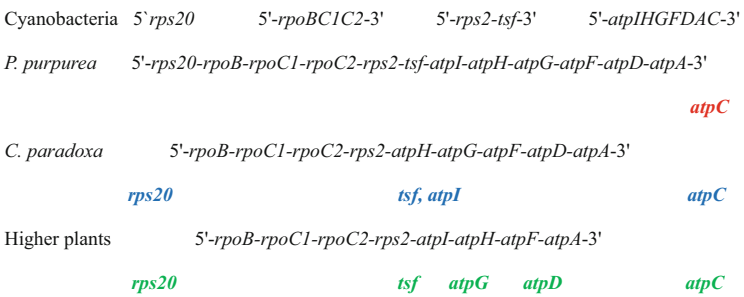


Fig. 2 Gene clusters of Archaeplastida diagnostic for plastid evolution. Genes indicated in *boldface* below the plastid operons were transferred to the nuclear genome in rhodophytes (*red*), glaucophytes (*blue*), and higher plants (*green*), respectively. The corresponding transcription units are widely separated on cyanobacterial genomes (*top*)

Outcome of Phylogenetic Analyses

A phylogenetic analysis based on 46 concatenated plastid genes from 10 species comprising glaucophytes, red algae, and green algae/plants clearly showed the monophyly of plastids, with glaucophytes on the first branch after the assumed single primary endosymbiotic event (Martin et al. 1998). The inclusion of only the then available *Synechocystis* 6803 as a reference cyanobacterium was criticized. However, a later repetition of such a tree including 15 cyanobacterial species yielded exactly the same result. A single event, i.e., monophyly of the kingdom "Plantae," should be demonstrated not only with respect to the endosymbiont but also with respect to the host. This task had to await extension of the nuclear gene databases from *Arabidopsis* to red algae and glaucophytes (for various reasons mitochondrial genes are not suitable for that purpose). A first attempt with a handful of nuclear genes yielded some support for a common origin of the red and green lineage (Moreira et al. 2000). With glaucophyte genes in the boat, in a concatenated analysis of 143 nuclear genes from 34 species good support for a sister group relationship of chlorophytes, rhodophytes, and glaucophytes was obtained—to the exclusion of all other eukaryotes (Rodríguez-Ezpeleta et al. 2005). Unlike in the plastid gene phylogeny, the branching order is not well resolved here and rhodophytes seem to branch first whereas other analyses using slightly different gene sets again place glaucophytes as the most ancient phototrophic eukaryotes known (Reyes-Prieto and Bhattacharya 2007a, b). Finally, the phylogenomic approach made possible with the completion of the *Cyanophora paradoxa* nuclear genome again corroborated the concept of a single primary endosymbiosis. Importantly, this result did not come from multigene trees but instead from the analysis of groups of genes involved in complex processes such as fermentation, plastid solute transport, and plastid protein import (Price et al. 2012). A more reliable determination of the branching order among Archaeplastida will be possible with the advent of additional genome sequences from glaucophytes and (mesophilic) rhodophytes (Chan et al. 2011). Remaining problems are the high percentage of LGT among

cyanobacteria (Dagan et al. 2008), a diverse and very ancient phylum where these processes take place since about 2.7 (Schirrmeister et al. 2013) to 2.4 (Rasmussen et al. 2008) billion years, i.e., before and after the primary endosymbiotic event. Additional complications are HGT among eukaryotes (Keeling and Palmer 2008; Chan et al. 2012) and potential phylogenetic artifacts (Stiller 2011). Consequently, a number of trees do not support a single event (Stiller 2007; Parfrey et al. 2010) or indicate a single event under the assumption of plastid loss in several eukaryotic lineages (Nozaki et al. 2007).

Insights from Comparing the Protein Import Machineries of Plastids

What other kind of "evidence" is available to help in answering this crucial question? There was the common opinion in the literature (McFadden and van Dooren 2004; Howe et al. 2008) that a decision must await more information about the respective protein import mechanisms of the different plastid types among the Archaeplastida: this is now available! The protein import apparatus of all primary plastids is homologous—a strong case for a single event (Steiner and Löffelhardt 2005)! This is demonstrated via heterologous in vitro import experiments with glaucophyte muroplasts and higher plant chloroplasts (Steiner et al. 2005) and by an analogous set of basic Toc and Tic translocon components (including those of cyanobacterial origin; Table 1) in all Archaeplastida lineages (Kalanon and McFadden 2008; Price et al. 2012; Sommer and Schleiff 2013). In particular, the pore-forming integral proteins allowing pre-protein translocation across both envelope membranes are common to all members of the Archaeplastida: Toc75/Omp85 (Sommer and Schleiff 2013) and Tic20 and/or Tic110 (Kovács-Bogdán et al. 2011), respectively. This does not mean that the whole import apparatus is identical: primordial plastids as muroplasts and rhodoplasts show a simpler, prototypic structure of the import apparatus, but their origin from a common ancestor, the protoplastid, is obvious (Steiner and Löffelhardt 2005; Table 1). With the appearance of the green lineage, a restructuring of the import apparatus became necessary to allow for efficient import of the most abundant soluble protein (Rubisco-SSU) and the most abundant membrane protein (LhcpII). Hence, some Toc components (e.g., Toc159 and likely Toc34) appeared only later in evolution in the derived chloroplasts, and also some Tic components (Tic214, Tic100, Tic56). The latter were recently shown to be components of a >1 MDa Tic complex in *Arabidopsis*, with Tic100 missing in green algae (Kikuchi et al. 2013). Tic214 had to be inserted into the chloroplast genome—a very unusual process in an organelle known for gene transfer and gene loss. In primordial plastids and in the few cases where the large Tic complex is missing in higher plants, (hetero)dimers of Tic20 family members are thought to be responsible for the general translocation channel in IM whereas the distinct Tic110 dimers might fulfill accessory or other functions

Table 1 Subunit composition of the plastid protein import apparatus of Archaeplastida

Subunit	Muroplast	Rhodoplast	Chloroplast	Function
Toc75[a]	+	+	+	Import pore (OM)
Toc34[b]	(+)	(+)	+	Receptor
Toc159[b]	–	–	–	Receptor
Tic20[a]	+	+	+	Import pore (IM)
Tic214[b,c]	–	–	+	Component of Tic20 complex[d]
Tic100[b]	–	–	+	Component of Tic20 complex[e]
Tic56[b]	–	–	+	Component of Tic20 complex[d]
Tic110[b]	+	+	+	Import pore (IM)?

[a]Provenience: cyanobacterial
[b]Provenience: eukaryotic(?)
[c]Plastome-encoded
[d]In Chloroplastida
[e]In higher plants

(Kikuchi et al. 2013). Convergent evolution cannot satisfactorily explain the observed situation. The preexisting mitochondrial import apparatus uses related targeting sequences, similar energy requirements, and also two collaborating translocons in OM and IM, but a largely different set of translocon components (Hewitt et al. 2013). *Paulinella chromatophora* is the result of a "primary endosymbiosis" insofar as a cyanobacterium and a heterotrophic protist were involved about 60 million years ago (Yoon et al. 2013). However, the endosymbionts don't (yet) represent primary plastids sensu stricto as the common protein import apparatus is lacking. Less than 30 documented and proven examples of gene transfer to the nucleus are known. Thus, protein import does not play a major role in the endosymbiont proteome and the import mechanism(s) still seem to be in the process of "evolutionary tinkering" (Novack and Grossman 2012). Therefore it is not appropriate to claim a "second primary endosymbiotic event" and use this as an argument against a single primary endosymbiosis (Parfrey et al. 2010). The *Paulinella* endosymbiosis happened more than 1,200 million years later between a member of a different clade of cyanobacteria and a member of a different protist lineage. Genome reduction of the cyanobacterial endosymbiont is 66 % as opposed to more than 95 % in the case of Archaeplastida. *P. chromatophora* isolates from Germany and Japan appear to be monophyletic (Yoon et al. 2009), indicating that successful endosymbiosis of a cyanobacterium and a protist is an extremely rare event. Further, the *Paulinella* endosymbionts harbor α-carboxysomes (Marin et al. 2007) like cyanobacteria, whereas pyrenoids represent the Rubisco microcompartments in primary plastids of algae (Price et al. 2012). The case of *Rhopalodia gibba* is different again (Adler et al. 2013). Here, the driving force is **not** autotrophy but nitrogen autonomy: the diatom host cell did already harbor a photosynthesizing plastid and the unpigmented cyanobacterial endosymbiont completely lost its capacity for photosynthesis. Due to the more recent origin (estimated to 25–35 million years ago), genome reduction is less pronounced (about 50 %) than for *P. chromatophora*. Taken together, the time frame for EGT in the amount of 95 %,

together with the establishment of a functional Toc/Tic translocase (allowing for a plastid proteome of >2,000), must exceed 100 million years by far. The vast complexity of the process alone renders the assumption of a single primary endosymbiotic event very likely.

The Cyanobacterial Ancestor

As plastid monophyly is generally accepted for some time, proposals for the likely cyanobacterial ancestor were made by several groups. (1) One of the favorites are filamentous nitrogen-fixing bacteria of subsection IV (Martin et al. 2002; Deusch et al. 2008). Their large genome sizes (compared to unicellular cyanobacteria) allow for EGT of several thousand genes which make up, e.g., 14 % of the *Arabidopsis* genome and are more similar in phylogenetic analyses to homologs from two filamentous cyanobacteria than to homologs from seven unicellular cyanobacteria. One other reason is the fact that extant *Nostoc* and *Anabaena* species are very prone to undergo (endo)symbioses (Adams et al. 2012) offering not only photoassimilates but also nitrogenous compounds to the host cells. Further, Omp85 of *Anabaena* was found closest to chloroplast Toc75 with respect to structure and pore-forming properties (Sommer and Schleiff 2013). A recent reevaluation taking into account that more than 60 % of cyanobacterial gene families have been affected by LGT again specifies modern diazotrophic filamentous cyanobacterial lineages (subsections IV and V) as having a gene collection most similar to that possessed by the plastid ancestor (Dagan et al. 2013). (2) Another candidate comes from among unicellular nitrogen-fixing cyanobacteria from the *Cyanothece* clade (subsection I; Ball 2013). These cyanobacteria store a starch-like reserve carbohydrate. The transition from (cyano)bacterial and host cell glycogen to the storage carbohydrate of phototrophic eukaryotes is easier explained in such a scenario. (3) A third group arrives at a different conclusion: an extensive phylogenomic analysis comprising 61 cyanobacterial taxa and 22 phototrophic eukaryotes does not support either proposal. Their data rather point towards an early emergence of primary plastids within cyanobacteria, prior to the diversification of most present-day cyanobacterial lineages (Criscuolo and Gribaldo 2011). Interestingly, more than 20 years ago, a limited data set of 16S ribosomal RNAs led to a similar view (Turner et al. 1999). However, the early emergence postulated might be attributable to long-branch attraction (Dagan et al. 2013) whereas a recent study confirmed an early appearance of filamentous cyanobacteria: subsection IV is considered the dominant cyanobacterial population between 1.5 and 1.2 billion years ago (Schirrmeister et al. 2013), which is about the time frame assigned to the primary endosymbiotic event (Parfrey et al. 2011).

Alternative Scenarios

The "shopping bag" model (Larkum et al. 2007; Howe et al. 2008) posits that repeated incorporations of cyanobacteria into phagocytosing protists with concomitant gene transfer to the nucleus paved the ground for a successful endosymbiosis later on. Recent data rather place nonphotosynthetic chlamydiae in that role: gene transfer during transitory infections of the host cell might have been the source of transporters, etc., important for the very early steps in primary endosymbiosis (Huang and Gogarten 2007; Facchinelli and Weber 2013; Ball 2013). Other authors prefer the explanation of LGT between cyanobacteria and many other bacterial lineages, including chlamydiae: the genome of the plastid ancestor entering into primary endosymbiosis already represented a genetic mosaic (Dagan et al. 2013).

Conclusions

On balance, primary endosymbiosis can be considered as a single event at the present state of knowledge. However, the shopping bag hypothesis certainly has its merits in secondary endosymbioses. Here, the "you are what you eat" principle (Doolittle 1998) might apply to an increased extent, at least when unicellular partners are considered. Multicellular animals though constantly feeding on algae show little or no tendency to incorporate prey-derived genes into their nuclear genomes, e.g., in the case of the sea slug *Elysia timida* (Wägele and Martin 2013).

References

Adams DG, Duggan PS, Jackson O (2012) Symbiotic interactions. In: Whitton B (ed) The ecology of cyanobacteria II: their diversity in space and time. Springer, Dordrecht, pp 593–664

Adl SM, Simpson AG, Farmer MA, Andersen RA, Anderson OR, Barta JR, Bowser SS, Brugerolle G, Fensome RA, Fredericq S, James TY, Karpov S, Kugrens P, Krug J, Lane CE, Lewis LA, Lodge J, Lynn DH, Mann DG, McCourt RM, Mendoza L, Moestrup O, Mozley-Standridge SE, Nerad TA, Shearer CA, Smirnov AV, Spiegel FW, Taylor MF (2005) The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. J Eukaryot Microbiol 52:399–451

Adler S, Trapp EM, Dede C, Maier UG, Zauner S (2013) *Rhopalodia gibba* – the first steps in the birth of a novel organelle? In: Löffelhardt W (ed) Endosymbiosis. Springer, Wien

Allen JF, de Paula WBM, Puthiyaveetil S, Nield J (2011) A structural phylogenetic map for chloroplast photosynthesis. Trends Plant Sci 16:645–655

Armbruster U, Pesaresi P, Pribil M, Hertle A, Leister D (2011) Update on chloroplast research: new tools, new topics, and new trends. Mol Plant 4:1–16

Baker A, Schatz G (1987) Sequences from a prokaryotic genome or the mouse dihydrofolate reductase gene can restore the import of a truncated precursor protein into yeast mitochondria. Proc Natl Acad Sci U S A 84:3117–3121

Ball SG (2013) Evolution of storage polysaccharide metabolism in Archaeplastida opens an unexpected window on the molecular mechanisms that drove plastid endosymbiosis. In: Löffelhardt W (ed) Endosymbiosis. Springer, New York, pp 111–134

Cavalier-Smith T (2000) Membrane heredity and early chloroplast evolution. Trends Plant Sci 5:174–182

Chan CX, Yang EC, Banerjee T, Yoon HS, Martone PT, Estevez JM, Bhattacharya D (2011) Red and green algal monophyly and extensive gene sharing found in a rich repertoire of red algal genes. Curr Biol 21:328–333

Chan CX, Bhattacharya D, Reyes-Prieto A (2012) Endosymbiotic and horizontal gene transfer in microbial eukaryotes: impacts on cell evolution and the tree of life. Mob Genet Elements 2:101–105

Criscuolo A, Gribaldo S (2011) Large-scale phylogenomic analyses indicate a deep origin of primary plastids within cyanobacteria. Mol Biol Evol 28:3019–3032

Dagan T, Artzy-Randrup Y, Martin W (2008) Modular networks and cumulative impact of lateral transfer in prokaryote genome evolution. Proc Natl Acad Sci U S A 105:10039–10044

Dagan T, Roettger M, Stucken K, Landan G, Koch R, Major P, Gould SB, Goremykin VV, Rippka R, Tandeau de Marsac N, Gugger M, Lockhart PJ, Allen JF, Brune I, Maus I, Pühler A, Martin WF (2013) Genomes of stigonematalean cyanobacteria (subsection V) and the evolution of oxygenic photosynthesis from prokaryotes to plastids. Genome Biol Evol 5:31–44

Deusch O, Landan G, Roettger M, Gruenheit N, Kowallik KV, Allen JF, Martin W, Dagan T (2008) Genes of cyanobacterial origin in plant nuclear genomes point to a heterocyst-forming plastid ancestor. Mol Biol Evol 25:748–761

Doolittle WF (1998) You are what you eat: a gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. Trends Genet 14:307–311

Facchinelli F, Weber APM (2013) Insertion of metabolite transporters into the endosymbiont membrane(s) as a prerequisite for primary endosymbiosis. In: Löffelhardt W (ed) Endosymbiosis. Springer, New York, pp 53–80

Fuentes I, Karcher D, Bock R (2012) Experimental reconstruction of the functional transfer of intron-containing plastid genes to the nucleus. Curr Biol 22:763–771

Gould SB, Waller RF, McFadden GI (2008) Plastid evolution. Annu Rev Plant Biol 59:491–517

Hagemann R (2007) The reception of the Schimper-Mereschkowski endosymbiont hypothesis on the origins of plastids between 1883 and 1960 –many negative, but a few relevant positive reactions. Ann Hist Phil Biol 12:41–59

Hewitt V, Lithgow T, Waller RF (2013) Modifications and innovations in the evolution of mitochondrial protein import pathways. In: Löffelhardt W (ed) Endosymbiosis. Springer, New York, pp 19–38

Howe CJ, Barbrook AC, Nisbet RER, Lockhart PJ, Larkum AWD (2008) The origin of plastids. Philos Trans R Soc Lond B Biol Sci 363:2675–2685

Huang JL, Gogarten JP (2007) Did an ancient chlamydial endosymbiosis facilitate the establishment of primary plastids? Genome Biol 8:R99

Huang CY, Ayliffe MA, Timmis JN (2003) Direct measurement of the transfer rate of chloroplast DNA into the nucleus. Nature 422:72–76

Kalanon M, McFadden GI (2008) The chloroplast protein translocation complexes of *Chlamydomonas reinhardtii*. A bioinformatic comparison of Toc and Tic components in plants, green algae, and red algae. Genetics 179:95–112

Keeling PJ (2010) The endosymbiotic origin, diversification and fate of plastids. Philos Trans R Soc Lond B Biol Sci 365:729–748

Keeling PJ, Palmer JD (2008) Horizontal gene transfer in eukaryotic evolution. Nat Rev Genet 9:605–618

Kikuchi S, Bédard J, Hirano M, Hirabayashi Y, Oishi M, Imai M, Takase M, Ide T, Nakai M (2013) Uncovering the protein translocon at the chloroplast inner envelope membrane. Science 339:571–574

Kovács-Bogdán E, Benz JP, Soll J, Bölter B (2011) Tic20 forms a channel independent of Tic110 in chloroplasts. BMC Plant Biol 11:133

Kowallik KV (1994) From endosymbionts to chloroplasts: evidence for a single prokaryotic/eukaryotic endocytobiosis. Endocytobiosis Cell Res 10:137–149
Lang BF (2013) Mitochondria and the origin of eukaryotes. In: Löffelhardt W (ed) Endosymbiosis. Springer, New York, pp 3–18
Larkum AWD, Lockhart PJ, Howe CJ (2007) Shopping for plastids. Trends Plant Sci 12:189–195
Leister D, Kleine T (2011) Role of intercompartmental DNA transfer in producing genetic diversity. Int Rev Cell Mol Biol 291:73–114
Lister DL, Bateman JM, Purton S, Howe CJ (2003) DNA transfer from chloroplast to nucleus is much rarer in *Chlamydomonas* than in tobacco. Gene 316:33–38
Lloyd AH, Timmis JN (2011) The origin and characterization of new nuclear genes originating from a cytoplasmic organellar genome. Mol Biol Evol 28:2019–2028
Löffelhardt W, Bohnert HJ, Bryant DA (1997) The complete sequence of the *Cyanophora paradoxa* cyanelle genome (Glaucocystophyceae). Plant Syst Evol 11:149–162
Margulis L (1981) Symbiosis in cell evolution. WH Freeman and Company, San Francisco, CA
Marin B, Nowack EC, Glöckner G, Melkonian M (2007) The ancestor of the *Paulinella* chromatophore obtained a carboxysomal operon by horizontal gene transfer from a *Nitrococcus*-like γ-proteobacterium. BMC Evol Biol 7:85
Martijn J, Ettema TJG (2013) From archaeon to eukaryote: the evolutionary dark ages of the eukaryotic cell. Biochem Soc Trans 41:451–457
Martin W (2003) Gene transfer from organelles to the nucleus: frequent and in big chunks. Proc Natl Acad Sci U S A 100:8612–8614
Martin W, Brinkmann H, Savona C, Cerff R (1993) Evidence for a chimaeric nature of nuclear genomes: eubacterial origin of eukaryotic glyceraldehyde-3-phosphate dehydrogenase genes. Proc Natl Acad Sci U S A 90:8692–8696
Martin W, Stoebe B, Goremykin V, Hansmann S, Hasegawa M, Kowallik KV (1998) Gene transfer to the nucleus and the evolution of chloroplasts. Nature 393:162–165
Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, Leister D, Stoebe B, Hasegawa M, Penny D (2002) Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proc Natl Acad Sci U S A 99:12246–12251
Matsuo M, Ito Y, Yamauchi R, Obokata J (2005) The rice nuclear genome continuously integrates, shuffles, and eliminates the chloroplast genome to cause chloroplast-nuclear DNA flux. Plant Cell 17:665–675
McFadden GI, van Dooren GG (2004) Evolution: red algal genome affirms a common origin of all plastids. Curr Biol 14:R514–R516
Mereschkowsky C (1905) Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. Biol Centralbl 25:593–604
Moreira D, Le Guyader H, Philippe H (2000) The origin of red algae and the evolution of chloroplasts. Nature 405:69–72
Novack EC, Grossman AR (2012) Trafficking of protein into the recently established photosynthetic organelles of *Paulinella chromatophora*. Proc Natl Acad Sci U S A 109:5340–5345
Nozaki H, Iseki M, Hasegawa M, Misawa K, Nakada T, Sasaki N, Watanabe M (2007) Phylogeny of primary photosynthetic eukaryotes as deduced from slowly evolving nuclear genes. Mol Biol Evol 24:1592–1595
Parfrey LW, Grant J, Tekle YI, Lasek-Nesselquist E, Morrison HG, Sogin ML, Patterson DJ, Katz LA (2010) Broadly sampled multigene analyses yield a well-resolved eukaryotic tree of life. Syst Biol 59:518–533
Parfrey LW, Lahr DJG, Knoll AH, Katz LA (2011) Estimating the timing of early eukaryotic diversification with multigene molecular clocks. Proc Natl Acad Sci U S A 108:13624–13629
Price DC, Chan CX, Yoon HS, Yang EC, Qiu H, Weber APM, Schwacke R, Gross J, Blouin NA, Lane C, Reyes-Prieto A, Durnford DG, Neilson JAD, Lang BF, Burger G, Steiner JM, Löffelhardt W, Meuser JE, Posewitz MC, Ball S, Arias MC, Henrissat B, Coutinho PM, Rensing SA, Symeonidi A, Doddapaneni H, Green BR, Rajah VD, Boore J, Bhattacharya D

(2012) *Cyanophora paradoxa* genome elucidates origin of photosynthesis in algae and plants. Science 335:843–847
Ran L, Larsson J, Vigil-Stenman T, Nylander JAA, Ininbergs K, Zheng W-W, Lapidus A, Lowry S, Haselkorn R, Bergman B (2010) Genome erosion in a nitrogen-fixing vertically transmitted endosymbiotic multicellular cyanobacterium. PLoS One 5:e11486
Rasmussen B, Fletcher IR, Brocks JJ, Kilburn MR (2008) Reassessing the first appearance of eukaryotes and cyanobacteria. Nature 455:1101–1104
Raven PH (1970) A multiple origin for plastids and mitochondria. Science 169:641–646
Reith M, Munholland J (1993) A high-resolution gene map of the chloroplast genome of the red alga *Porphyra purpurea*. Plant Cell 5:465–475
Reyes-Prieto A, Bhattacharya D (2007a) Phylogeny of nuclear-encoded plastid-targeted proteins supports an early divergence of glaucophytes within Plantae. Mol Biol Evol 24:2358–2361
Reyes-Prieto A, Bhattacharya D (2007b) Phylogeny of calvin cycle enzymes supports plantae monophyly. Mol Phylogenet Evol 45:384–391
Ris H, Plaut W (1962) Ultrastructure of DNA-containing areas in the chloroplast of *Chlamydomonas*. J Cell Biol 13:383–391
Rodríguez-Ezpeleta N, Brinkmann H, Burey SC, Roure B, Burger G, Löffelhardt W, Bohnert HJ, Philippe H, Lang BF (2005) Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes. Curr Biol 15:1325–1330
Rousseau-Gueutin M, Ayliffe MA, Timmis JN (2011) Conservation of plastid sequences in the plant nuclear genome for millions of years facilitates endosymbiotic evolution. Plant Physiol 157:2181–2193
Sagan L (1967) On the origin of mitosing cells. J Theor Biol 14:225–275
Schirrmeister BE, de Vosb JM, Antonellic A, Bagheria HC (2013) Evolution of multicellularity coincided with increased diversification of cyanobacteria and the Great Oxidation Event. Proc Natl Acad Sci U S A 110:1791–1796
Smith DR, Crosby K, Lee RW (2011) Correlation between nuclear plastid DNA abundance and plastid number supports the limited transfer window hypothesis. Genome Biol Evol 3:365–371
Sommer MS, Schleiff E (2013) Evolution of the protein translocons of the chloroplast envelope. In: Löffelhardt W (ed) Endosymbiosis. Springer, New York, pp 81–110
Stegemann S, Bock R (2006) Experimental reconstruction of functional gene transfer from the tobacco plastid genome to the nucleus. Plant Cell 18:2869–2878
Steiner JM, Löffelhardt W (2005) Protein translocation into and within cyanelles. Mol Membr Biol 22:123–132
Steiner JM, Yusa F, Pompe JA, Löffelhardt W (2005) Homologous protein import machineries in chloroplasts and cyanelles. Plant J 44:646–652
Stiller JW (2007) Plastid endosymbiosis, genome evolution and the origin of green plants. Trends Plant Sci 12:391–396
Stiller JW (2011) Experimental design and statistical rigor in phylogenomics of horizontal and endosymbiotic gene transfer. BMC Evol Biol 11:259
Stiller JW, Reel DC, Johnson JC (2003) A single origin of plastids revisited: convergent evolution in organellar genome content. J Phycol 39:95–105
Stoebe B, Kowallik KV (1999) Gene-cluster analysis in chloroplast genomics. Trends Genet 15:344–347
Stoebe B, Maier UG (2002) One, two, three: nature's tool box for building plastids. Protoplasma 219:123–130
Thiergart T, Landan G, Schenk M, Dagan T, Martin WF (2011) An evolutionary network of genes present in the eukaryote common ancestor polls genomes on eukaryotic and mitochondrial origin. Genome Biol Evol 4:466–485
Timmis JN, Ayliffe MA, Huang CY, Martin W (2004) Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nat Rev Genet 5:123–135

Turner S, Pryer KM, Miao VPW, Palmer JD (1999) Investigating deep phylogenetic relation-ships among cyanobacteria and plastids by small subunit rRNA sequence analysis. J Eukaryot Microbiol 46:327–338

Wägele H, Martin WF (2013) Endosymbioses in sacoglossan seaslugs: plastid-bearing animals that keep photosynthetic organelles without borrowing genes. In: Löffelhardt W (ed) Endosymbiosis. Springer, New York, pp 291–324

Wallin IE (1925) On the nature of mitochondria. IX. Demonstration of the bacterial nature of mitochondria. Am J Anat 35:131–149

Williams TA, Foster PG, Nye TMW, Cox CJ, Embley TM (2012) A congruent phylogenomic signal places eukaryotes within the Archaea. Proc R Soc B 279:4870–4879

Yoon HS, Nakayama T, Reyes-Prieto A, Andersen RA, Boo SM, Ishida KI, Bhattacharya D (2009) A single origin of the photosynthetic organelle in different *Paulinella* lineages. BMC Evol Biol 9:98

Yoon HS, Yang EC, Qiu H, Bhattacharya D (2013) Photosynthetic *Paulinella* –recapitulation of primary plastid establishment. In: Löffelhardt W (ed) Endosymbiosis. Springer, New York, pp 151–166

Zybailov B, Rutschow H, Friso G, Rudella A, Emanuelsson O, Sun Q, van Wijk KW (2008) Sorting signals, N-terminal modifications and abundance of the chloroplast proteome. PLoS One 3:e1994

Insertion of Metabolite Transporters into the Endosymbiont Membrane(s) as a Prerequisite for Primary Endosymbiosis

Fabio Facchinelli and Andreas P.M. Weber

Abstract Eukaryotes acquired the ability to perform photosynthesis by capturing and stably integrating a photoautotrophic prokaryote. This event, referred to as primary endosymbiosis, occurred only once in the ancestral protoalga, giving rise to the *Archaeplastida* comprising three major photoautotrophic lineages: the glaucophytes, the red algae, and the green algae.

One crucial step for the success of the endosymbiosis must have been the integration and coordination of the metabolism of the host and the endosymbiont. Metabolic integration requires traffic of metabolites across the envelope membrane, which represents the specificity barrier separating the cyanobiont from the host cell cytoplasm. Insertion of translocators into the endosymbiont's envelope was necessary to ensure a controlled exchange of molecules and to take full advantage of the newly acquired metabolic entity. Based on genome sequence data, phylogenetic analyses revealed that the major contribution in establishing a connection between the two partners was provided by the host cell in order to rapidly take control over the endosymbiont, with a minor contribution coming from the cyanobiont and from a third chlamydial source. With this chapter we provide an update on recent findings in elucidating the repertoire of plastidic transporters with a focus on their evolutionary history.

Contents

Introduction 54
Transport of Metabolites 59
Carbon Metabolism 59
Nitrogen Metabolism 66

F. Facchinelli • A.P.M. Weber (✉)
Institut für Biochemie der Pflanzen, Heinrich-Heine Universität Düsseldorf, Universitätstrasse 1, 40225 Düsseldorf, Germany
e-mail: Andreas.Weber@uni-duesseldorf.de

W. Löffelhardt (ed.), *Endosymbiosis*, DOI 10.1007/978-3-7091-1303-5_4,
© Springer-Verlag Wien 2014

Evolutionary Origin of the Plastidic Translocators 70
Intracellular Parasites as Drivers for the Metabolic Integration 71
Conclusions 72
References 73

Introduction

Life is characterized by a nonequilibrium redox chemistry where a continuous flux of energy is needed for organisms to persist and proliferate. This energy is stored in the chemical bonds of the building blocks making up the organic macromolecules. The ultimate source of energy is almost exclusively derived from the sun. Photoautotrophic organisms performing oxygenic photosynthesis trap the energy contained in the photons and use it to strip electrons and protons from water. These electrons are then used to assimilate inorganic carbon (but also nitrogen, phosphate, and sulfur) to build up cellular components (Falkowski and Godfrey 2008; Hohmann-Marriott and Blankenship 2011).

Long before oxygenic photosynthesis arose, the biochemistry of the organisms did not rely on the oxygen–water cycle but instead used stronger electron donors, such as hydrogen, hydrogen sulfide, and ferrous iron. Geological records date back the first photosynthetic organisms using H_2 to 3.8 billion years ago, soon after the end of the late heavy bombardment by asteroids (Olson 2006). Today these anoxygenic photosynthetic bacteria account for <0.1 % of total primary productivity (Raven 2009). Anoxygenic photosynthesis requires only one of the two types of photochemical reaction centers. The purple photosynthetic bacteria perform a cyclic electron transport around a quinone-containing type II reaction center resembling the photosystem II of oxygenic photosynthetic organisms that oxidize water and reduce plastoquinone. The cyclic electron transfer is coupled to the generation of a proton motive force across a membrane that is used to generate ATP.

Green-sulfur bacteria possess instead a FeS-based reaction center (type I) similar to the photosystem I of oxygenic photosynthesizers, which drives linear electron transfer from hydrogen sulfide or iron to ferredoxin and ultimately $NADP^+$ (Olson and Blankenship 2004; Bryant and Frigaard 2006). Despite their low sequence identity, the two reaction centers are homologues, meaning that they descend from a common ancestor, as indicated by their structural overlay (Sadekar et al. 2006). Phylogenetic and structural analyses point to an early duplication and divergence of the reaction center ancestor which subsequently resulted in the heterodimeric type I and type II reaction centers (Allen 2005). The current hypothesis states that the protocyanobacterium possessed both types of reaction centers but used only one of them at a time. Redox regulation of gene expression determines whether type I or type II reaction center genes are expressed: when hydrogen sulfide is present, the bacteria switch to type I reaction centers to produce reduction equivalents for making up organic matter. When the environment changes towards less reducing

conditions, they rely on the type II reaction center to produce ATP (Allen 2005; Allen and Martin 2007). The latter state has an intrinsic drawback: the electron circuit can be over-reduced by electrons from the environment. One such electron source could have been manganese, which is present in high amounts in the oceans and acts as antioxidant by absorbing ultraviolet radiation, thereby pushing electrons into the photooxidized carriers of photosystem II. To overcome this impasse, all what bacteria needed to do was to disable the redox switch preventing the concomitant expression of both photosystems. The simultaneous expression of photosystem I and II would offer a way out for the electrons clogged in the photosystem II, hence providing the protocyanobacterium with the advantages of both models of photosynthesis: ATP synthesis and reduction of electron carriers (Allen 2005). The last step to accomplish for evolving oxygenic photosynthesis would be to trap the electron source, manganese, close to the reaction center of photosystem II. Modern-day oxygenic photosynthesizers possess an oxygen-evolving cluster consisting of four manganese atoms held together with five oxygen atoms and a single calcium atom. The structure of this cluster was recently resolved at a resolution of 1.9 Å (Umena et al. 2011). The structure of this complex is remarkably similar to the manganese oxide minerals such as hollandite found in the ocean depths where water splitting is thought to have arisen and suggests that one of the core reactions of photosynthesis is the product of a mineral catalyst (Russell 2006). The manganese cluster, when fully oxidized by chlorophyll *a* reaction centers, takes back four electrons from two water molecules, thereby releasing protons and molecular oxygen. Photosynthesis is now oxygenic and no more dependent on the availability of H_2S and ferrous iron, but instead runs at the expense of water, which, being available at a concentration of 56 M, can be seen as an infinite source of electrons.

The by-product of photosynthesis, oxygen, is maybe the most valuable waste product on earth: without oxygen there would be no ozone protecting the earth's surface from the deleterious effects of ultraviolet rays. UV light slowly splits water into oxygen and hydrogen, the former oxidizing the iron in the crust's rocks and never accumulating in the air, and the latter, the lightest gas, evades from the atmosphere. Without oxygen accumulation the oceans would have bled into the space, the same fate faced by the oceans of Venus (Lane 2009). Along with its role in building up the atmosphere shield, oxygen enabled aerobic respiration to occur and thus a more efficient way to obtain energy from reduced molecules. Oxygen is a strong oxidizing agent with a redox potential of the couple $H_2O/1/2\ O_2$ of +0.815 V, enabling the complete oxidation of reduced carbon compounds to carbon dioxide and water. The energetic efficiency of oxygenic respiration enabled the evolution of multicellular forms of life and permitted the establishment of longer food chains (Lane 2009). Without the rise of oxygen concentration in the earth's atmosphere, our planet would probably still be dominated by bacterial life forms.

Oxygenic photosynthesis makes use of a complex machinery consisting of more than 100 genes. Among the prokaryotes, only cyanobacteria evolved the ability to perform oxygenic photosynthesis, indicating that the lateral transfer is precluded in the otherwise metabolically versatile bacteria, due to the resistance of the core

photosynthetic apparatus to horizontal gene transfer (HGT). This is consistent with the complexity hypothesis, which states that genes coding for large complex systems that have more macromolecular interactions are less subject to HGT than genes coding for small assemblies of a few gene products (Jain et al. 1999; Shi and Falkowski 2008). The complexity of oxygenic photosynthetic machineries together with their distribution across different operons makes it difficult to transfer it to non-photosynthetic prokaryotes.

Rather, the transfer of the photosynthetic machinery could only be accomplished by the wholesale conveyance of the cyanobacterium to a heterotrophic cell. This event, referred to as primary endosymbiosis, involves the captive retention of a cyanobacterium by a primitive mitochondriate eukaryote and states the birth of the photosynthetic eukaryotic lineage. The primary endosymbiotic event occurred between 1 and 1.5 billion years ago, leading to the ancestor of the plastid-containing modern algae (Yoon et al. 2004; Rodríguez-Ezpeleta et al. 2005). The resemblance between cyanobacteria and plastids was already noted by the botanist Andreas Schimper at the end of the nineteenth century, and few decades later Mereschkowski formalized the hypothesis that the chloroplasts derive from cyanobacteria (Schimper 1883; Mereschkowsky 1905). Later on in the 1970s, the theory of endosymbiosis became widely accepted thanks to Lynn Margulis who proposed that mitochondria and plastids (and in a first moment also cilia and flagella) all derive from a bacterial endosymbiosis (Sagan 1967). Genome sequences and phylogenies of plastid-encoded genes and of plastid-targeted nuclear genes support a single, ancient origin of the photoautotrophic eukaryotes (Wolfe et al. 1994; Martin et al. 1998; Moreira et al. 2000). Once the cyanobacterial endosymbiont was established, the three major lineages of Archaeplastida diverged: the glaucophytes, the red algae, and the green algae (Adl et al. 2005). The plastids originating from primary endosymbiosis are characterized by the presence of two membrane layers separating them from the cytosol and showing a chimeric composition with both bacterial signatures (the presence of galactolipids of cyanobacterial origin and of β-barrel proteins) and host features such as the replacement of the lipopolysaccharide with phosphatidylcholine shuttled from the ER to the outer leaflet of the outer membrane (Cavalier-Smith 1982, 2000; Schleiff et al. 2003).

The glaucophytes were likely the first lineage diverging after plastid establishment (Martin et al. 1998). Glaucophytes comprise a small group of chlorophyll *a*-containing unicellular freshwater algae. Their basal position to the red and green algae is also reflected by their plastids which most closely resemble the cyanobacterial progenitor. The plastid, also referred to as muroplast, still retains a peptidoglycan wall between the two membranes, a relic of its bacterial ancestor and a strong proof for the endosymbiotic theory (Steiner and Löffelhardt 2002; Sato et al. 2009). In addition, thylakoid membranes of the glaucophytes are decorated with phycobilisome light-harvesting antennae very similar to those found in cyanobacteria.

Red algae (rhodophytes) also contain phycobilisomes similar to those found in cyanobacteria. Moreover, they possess light-harvesting complex I (LHCI)-type

antenna proteins homologous to those of plants. In addition, PS I-associated phycobilisome substructures (consisting of a rod only) were reported for *Cyanidioschyzon merolae*. This suggests that three different types of antenna systems exist in the rhodoplast that can interact with photosystem I, thus representing an intermediate state between the cyanobacterium and the chloroplast (Wolfe et al. 1994; Vanselow et al. 2009; Busch et al. 2010).

Green algae and land plants (chlorophyta) lack the extrinsic phycobilisomes which were substituted by the membrane-embedded, chlorophyll *a*/*b*-containing LHCs associated with both photosystems, along with a diverse set of accessory pigments devoted to light capture and photoprotection.

While the evolutionary history of the primary endosymbiotic event is generally accepted, the picture of plastid evolution gets complicated by eukaryotic endosymbioses where heterotrophic eukaryotes acquired the ability to perform photosynthesis by capturing a green or a red alga (Cavalier-Smith 2003; Bhattacharya et al. 2004).

Acquisition of phototrophy through secondary endosymbiosis was the preferred path by which most algal lineages acquired their plastids. These algae not only contributed to the eukaryotic diversity by giving rise to a plethora of new species, but also represent the dominating primary producers of the aquatic environment.

Secondary endosymbiosis took place by the engulfment of a green or red alga, which then progressively degenerated to a plastid which retained the footprint of this event in form of one or two extra membranes, and, in few cases, a remnant of the algal nucleus, the nucleomorph (Gould et al. 2008; Keeling 2010). How many times secondary endosymbioses have happened is still a matter of debate. The most simplistic scenario involves only two secondary endosymbioses: one in the green lineage leading to the Cabozoa, which include the euglenophytes and the chlorarachniophytes, and one in the red lineage leading to the Chromalveolata (comprising the nucleomorph-containing cryptophytes, the haptophytes, the plastid-bearing heterokontophytes such as diatoms, and the alveolates; Cavalier-Smith 1999). Analyses of plastid genomes now refute the Cabozoa hypothesis and point to two independent acquisitions, whereas for the Chromalveolata hypothesis the picture is complicated by the far greater number of lineages involved and a single origin of their plastids is still under debate (for details, see Gould et al. 2008; Keeling 2010 and references therein). Despite the different scenarios which can explain the distribution of plastids in the Plantae kingdom, it appears that the establishment of this first association between the cyanobacterium and the eukaryotic host may have been a more difficult challenge than the subsequent symbioses involving the uptake of already integrated (with respect to plastids) algae by diverse eukaryotic hosts.

Modern cyanobacterial genomes code for a few thousands of proteins, while plastid genomes rarely exceed a coding capacity of 100 gene products. Primary plastids have undergone a massive reduction of their plastome coding capacity, with the majority of their genes transferred to the host nucleus early after the endosymbiotic association, as indicated by the shared complement of plastid-encoded genes (Martin et al. 1998; Timmis et al. 2004). The relocation of the

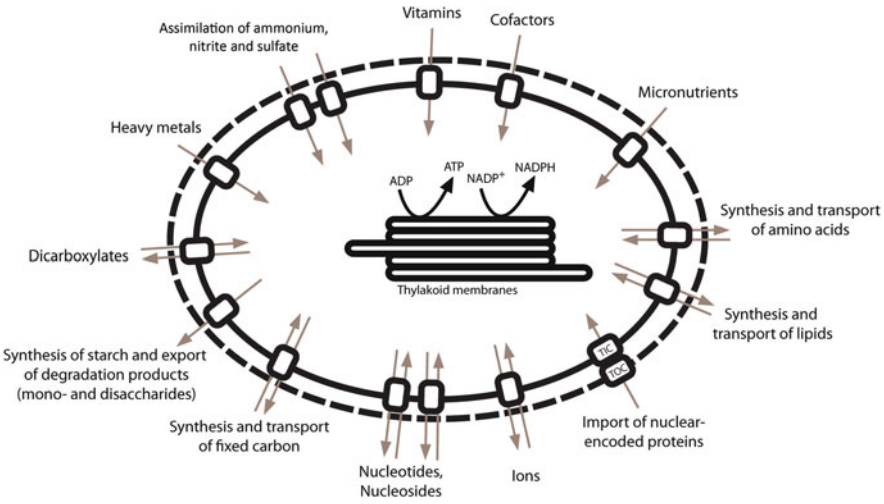


Fig. 1 Schematic representation of the major plastidic metabolic pathways and the transporters localized to the envelope membrane responsible for the shuttling of the intermediates and products. The presence of different pathways and transporters varies according to the developmental stage, the tissue, and the evolutionary history of the algal group

gross of the plastid genome to the nucleus by endosymbiotic gene transfer (EGT) on one side cemented the interdependence of the cyanobacterium from the host cell thus turning the initial association into an obligate one and on the other side pushed for the development of a mechanism to redirect the gene products back to the original compartment. This was achieved by the TOC–TIC import machinery responsible for the translocation of plastid-targeted proteins which harbor a specific topogenic signal at their N-terminus. The translocon most likely evolved before the split of the Archaeplastida, since elements are present which are shared among the three major autotrophic lineages (Steiner and Löffelhardt 2002; Gross and Bhattacharya 2009). The hallmark of the plastid-containing eukaryotes is the ability to rely on the reduction of carbon dioxide for their energetic and biosynthetic requirements by performing photosynthesis. In addition to that, the cyanobacterium introduced into the protoalga a diverse set of anabolic and catabolic pathways which reflect the metabolic versatility typical of the prokaryotic kingdom, ranging from nitrogen and sulfur assimilation to the biosynthesis of fatty acids, amino acids, hormones, and a plethora of secondary compounds (Weber and Flügge 2002; Weber and Fischer 2007; Gould et al. 2008).

The exchange of precursors, intermediates, and end products of the newly acquired metabolic pathways necessitates a high flux of metabolites across the plastid envelope membranes. This is accomplished by various transporter proteins located in the innermost membrane of the plastid ensuring a controlled exchange of metabolites between the different compartments (Fig. 1; Facchinelli and Weber 2011; Weber and Linka 2011).

This chapter focuses on the importance of such transporters for the establishment of a reliable connection between the host and the early endosymbiont and the essential function of metabolite transporters for the coordination of the two newly integrated metabolic entities. We will point to the different contributions both from the cyanobacterial endosymbiont and from the host cell as sources for the metabolite transporters. We will further extend the discussion to a third, minor but apparently essential partner, which recently turned out as having contributed to the early events during the establishment of primary endosymbiosis.

Transport of Metabolites

Plastids are multifaceted organelles able not only to convert the solar energy into stable chemical bonds through the process of photosynthesis, but they are also responsible for the storage of a wide variety of products and for the synthesis and transformation of key cellular compounds. Among the metabolic routes located in the plastid, the most prominent ones include the biosynthesis of chlorophylls, carotenoids, purines, pyrimidines, fatty acids, amino acids, and a range of secondary metabolites. Plastids also reduce the important inorganic ions nitrite and sulfate, both imported from the cytosol.

The metabolite traffic between the plastid stroma and the cytosol is predominantly facilitated by antiporters embedded in the inner envelope membrane. The outer envelope has long been thought as not being involved in the discrimination of the compounds shuttled between the two compartments, and that the transit across this lipid bilayer is only limited by size (Flügge 2000). The discovery of low-affinity, high-specific porin-like channels in the outer envelope suggests an additional layer of control of the metabolic traffic across the plastid boundary (Soll et al. 2000; Duy et al. 2007). The selectivity barrier is however given by the inner envelope which hosts several transporters with distinct substrate specificities (Flügge 1998). A detailed description of these transporters is provided in several recent reviews (Linka and Weber 2010; Facchinelli and Weber 2011; Weber and Linka 2011).

Carbon Metabolism

The main function of autotrophic plastids is the net assimilation of carbon dioxide through photosynthesis. CO_2 enters the Calvin–Benson cycle by reacting with a five-carbon sugar, ribulose-1,5-bisphosphate (RuBP). This step is catalyzed by the RuBP carboxylase/oxygenase (Rubisco), one of the three enzymes unique to the Calvin–Benson cycle. The carboxylation of the C_5 sugar produces a C_6 intermediate which is immediately cleaved into two molecules of 3-phosphoglyceric acid (3-PGA). Following the carboxylation, a two-step reductive phase converts the

3-PGA into the triose phosphate glyceraldehyde-3-phosphate (GAP). This step requires energy and reducing equivalents, in the form of ATP and NADPH produced through photosynthetic electron transport. Out of six GAP molecules synthesized during the carboxylation/reduction phases, five are used to regenerate the CO_2 acceptor RuBP, while one GAP can be withdrawn from the Calvin–Benson cycle and represents the net gain of the photosynthetic pathway. The assimilated carbon can have different destinations according to the cellular needs, its allocation being controlled at several levels. In the Viridiplantae, reduced carbon can be stored in the form of insoluble starch inside the plastid or soluble sucrose for long-distance transport or directed to structural carbohydrate biosynthesis (cellulose). Triose phosphates are also the substrate of a wide range of anabolic pathways residing both inside and outside the chloroplast.

Transporters Involved in Primary Carbon Metabolism

Triose phosphates (TP) are exported by a member of the plastidic phosphate translocators (pPTs), namely the triose phosphate/phosphate translocator (TPT). In the model plant *Arabidopsis thaliana*, the pPT family comprises four members belonging to the drug/metabolite transporters (DMT) superfamily. pPTs are more closely related to the nucleotide sugar transporters (NSTs) than to other families of the DMT superfamily (Knappe et al. 2003). These antiporters have partially overlapping but distinct substrate specificities and catalyze a strict counter-exchange of phosphorylated C_3-, C_5-, and C_6-compounds for inorganic phosphate (P_i). The homo-exchange of *ortho*-phosphate versus phosphorylated sugars guarantees the homeostasis of the phosphate content inside the stroma to satisfy the needs of the light reactions where phosphate is one of the substrates of photophosphorylation (Weber et al. 2005).

As outlined above, triose phosphates leave the plastid through the TPT. TPT represents in plants the route for assimilated carbon export during the day and it accepts TPs (dihydroxyacetone phosphate (DHAP) and GAP) as well as 3-PGA as substrates (Fliege et al. 1978; Flügge and Heldt 1984; Flügge et al. 2003). Besides its role in allocating sugars between the stroma and the cytosol, the TPT can also act as a shuttle for reduction equivalents by exporting triose phosphates to the cytosol where they are converted to 3-PGA by the action of the GAP dehydrogenase, thereby reducing one molecule of NAD^+ to NADH. The 3-PGA crosses the plastid envelope through the same transporter and can be further metabolized by the Calvin–Benson cycle (Flügge and Heldt 1984).

For the red alga *Galdieria sulphuraria*, liposome uptake experiments with reconstituted phosphate translocators isolated from total membranes demonstrated the activity for transport of TP, but not 3-PGA (Weber et al. 2004). Heterologous expression of the TPT homologue from *G. sulphuraria* confirmed its narrower substrate specificity, i.e., for TPs only. In addition, the red algal transporter showed a twofold higher affinity for its substrate, in comparison to the plant TPT (Linka et al. 2008). While green algae and land plants store assimilated carbon in form of

insoluble starch granules inside the chloroplast stroma, red algae synthesize a particular polymer, called floridean starch, starting from UDP-glucose moieties and resembling amylopectin in its structure (Viola et al. 2001; Ball and Morell 2003). In red algae, UDP-glucose is also utilized to synthesize floridoside, the main soluble pool of storage carbon in analogy with sucrose in higher plants. Rhodophytes thus allocate photoassimilates exclusively to the cytosol. Red algae, under active photosynthetic conditions, require TPT activity solely for carbon export, in order to cope with the high demand of fixed carbon in the cytosol.

Among the glaucophytes, *Cyanophora paradoxa* is considered the model organism of reference (Löffelhardt et al. 1997). Similarly to the red algae, glaucophytes also store the starch pool (relatively amylose rich—a parallel to chlorophytes) in the cytosol, using UDP-glucose and, potentially, also ADP-glucose (Plancke et al. 2008). Transport experiments with isolated cyanelles could demonstrate that phosphate uptake is inhibited by DHAP, 3-PGA, and glucose-6-phosphate (Glc6-P), pointing to the presence of a TPT as in the other two Archaeplastida lineages and thus providing the evidence of the evolution of the phosphate translocator in the algal ancestor (Schlichting and Bothe 1993; Schlichting et al. 1994). The recently sequenced nuclear genome of *C. paradoxa* (Price et al. 2012), however, suggests another picture regarding the mechanism of carbon export from the cyanelles of the glaucophyte algae, as outlined below.

Starch Metabolism

A crucial point for a successful establishment of a symbiotic relationship between the cyanobacterium and the host cell would have been the connection of the metabolism of both organisms. This provokes the question as to how the mitochondriate heterotrophic cell would have profited from the cyanobacterium's ability to synthesize carbohydrates through photosynthesis. The current hypothesis states that the cyanobacterial ancestor involved in the primary endosymbiosis was related to modern-day subgroup V cyanobacteria (Deschamps et al. 2008a). These are single-cell diazotrophic cyanobacteria that lack the ability to form filamentous colonies which develop specialized structures (heterocysts) devoted to nitrogen fixation where the nitrogenase is sheltered from oxygen damage. This group of bacteria temporally separate oxygenic photosynthesis and nitrogen fixation through circadian clock regulation (Schneegurt et al. 1994). Group V cyanobacteria are the sole group of bacteria able to synthesize starch, the form of carbon storage also found in Archaeplastida and their descendants originated by secondary endosymbiosis. It is hypothesized that the ability to synthesize starch arose in the unicellular diazotrophic cyanobacteria because of its resistance to hydrosoluble degrading enzymes and because, unlike glycogen, starch is osmotically inert and not subjected to size limitation, thereby providing a form of photoassimilate polymerization suitable for the long-term storage. Accumulation of starch during the day could thus fuel the energy-demanding process of nitrogen assimilation during the night (Deschamps et al. 2008a).

All three lineages derived by primary endosymbiosis seem to have gained the ability to synthesize starch at a very early stage. The pathway of starch synthesis consists of a mosaic of genes of cyanobacterial and host origin, but the absence of starch from other eukaryotes excluding algae and plants suggests the presence of glycogen rather than starch in the ancestral host cell (Deschamps et al. 2008b, c). The eukaryotic and prokaryotic pathways of storage glucan synthesis differ in the nature of the activated sugar used for polymerization: while heterotrophic eukaryotes use UDP-glucose, bacteria utilize ADP-glucose. The original path for starch biosynthesis is thought to have occurred in the cytosol of the protoalga, as still is the case for the rhodophytes and the glaucophytes (Deschamps et al. 2008a). According to this scenario, the cyanobiont lost storage polysaccharide metabolism early after the primary endosymbiotic event. The current hypothesis states that at the very beginning of the endosymbiosis, a dual pathway for glycogen biosynthesis existed in the cytosol of the protoalga: the original eukaryotic pathway based on UDP-glucose, and one based on ADP-glucose exported from the endosymbiont and polymerized by a cyanobacterial starch synthase representing a very early EGT to the nucleus (Ball et al. 2011).

Metabolic symbiosis was achieved by the transfer of genes involved in starch biosynthesis from the organelle to the host nucleus and their expression, under host promoters, in the cytosol. This implies the presence in the cytosol of ADP-glucose coming from the cyanobiont through some sort of transporter able to export the sugar nucleotide. Interestingly, the pPTs, which are responsible for carbon partitioning in the Plantae, derive from a NST of host origin (Weber et al. 2006). Thus, the generation of metabolic flux in the ancient endosymbiont likely had required just a single gene transfer (an ADP-glucose-dependent starch synthase) accompanied by the targeting of an existing transporter to the cyanobacterial envelope membrane.

This step does not require any targeting of the gene products back to the endosymbiont and may thus predate the origin of the plastid protein import machinery. Transfer of a minimal set of genes (isoamylase and disproportionating enzyme) would later on have been enough to establish a cytosolic starch biosynthetic pathway necessary to store the large amounts of sugars made available by the endosymbiont (Deschamps et al. 2008a).

Evolutionary History of the Phosphate Translocators

Phylogenetic analyses revealed that the pPT family of translocators is monophyletic and evolved from an existing nucleotide sugar transporter (NST) of the endomembrane system of the host cell (Weber et al. 2006). Available sequences from red algae, green algae, and plants show that the different members of the pPT family have homologues in both lineages, indicating an early origin of these translocators in the common ancestor of the Archaeplastida (Weber et al. 2006). Members of the NST family localize to the ER and Golgi apparatus of eukaryotic cells and are involved in the uptake of activated monosaccharides for glycosylation

reactions. The relatedness of the pPTs with the NSTs is intriguing since this could represent the missing link between the proposed ancestral metabolic connection in the protoalga and the modern carbon exporters. NSTs, however, do not physiologically transport ADP-glucose since this metabolite is exclusively found in prokaryotes. Interestingly, purine nucleotide glycosyl transporters are present within the subfamily M of NSTs, which forms a clade with the pPTs. Indeed, transport experiments with two members of the NSTs responsible for the transport of GDP-sugars displayed an innate ability to transport ADP-glucose (Colleoni et al. 2010). The transport of the substrate GDP-mannose by the two transporters tested (Vrg4p from yeast and GONST1 from *Arabidopsis*) was efficiently competed by physiological concentrations of AMP. Moreover, the *Arabidopsis* homologue was able to exchange ADP-glucose for AMP, although the affinity for ADP-glucose was rather low compared with GDP-mannose (250-fold higher affinity for GDP-mannose). The authors argue that in the cytosol the concentrations of GDP-mannose and other GDP-sugars are very low (below 50 μM) compared to the concentration of ADP and AMP, kept high by the polymerization of starch, and that the amount of AMP would have been enough to outcompete GDP-mannose and favor the entry of AMP into the endosymbiont (Colleoni et al. 2010). As previously stated, the cyanobacterium likely lost very early the ability to synthesize polysaccharides, behaving like a mutant impaired in the utilization of ADP-glucose. This would have enabled the accumulation of ADP-glucose in amounts compatible with the kinetic properties of the NST, permitting its efflux. In this respect, a cytotoxic effect following the insertion of the transporter was improbable since only carbon destined for storage was withdrawn from the endosymbiont (Colleoni et al. 2010; Ball et al. 2011; Linka and Weber 2012).

The first metabolic link between the two symbionts thus implicates the insertion of an ADP-glucose translocator into the membrane of the early plastid. Later on, the ancient ADP-glucose translocator was substituted by a member of the pPT gene family. The reason for this substitution was seen in the evolution of the protein import machinery: as the metabolic connections between the two partners were tightened by the targeting of gene products to the organelle, gene duplications and divergence would have enabled a more optimized integration of the metabolic pathways. This then resulted in the diverse pPTs which are responsible for the shuttling of reduced carbon compounds across the plastid envelope of extant plastids. The NST and its derivatives, the pPTs, may have coexisted as long as starch was synthesized from ADP-glucose in the cytosol. Loss of the ADP-glucose-dependent pathway in the rhodophytes and glaucophytes and its return to the plastid in the Viridiplantae posed no need for retaining the NST and its function was taken over by the pPTs (Colleoni et al. 2010). Besides the already mentioned TPT responsible for the daily carbon export, the pPT family members evolved novel transporter activities exemplified by the GPT (Glc6-P/phosphate translocator), the XPT (xylulose-5-phosphate/phosphate translocator), and the PPT (phospho*enol*pyruvate/phosphate translocator). This set of transporters perfectly fulfills the metabolic requirements of higher plants, and differences between these and red algal

pPTs reflect the different needs of carbon partitioning between their subcellular compartments, as exemplified for the TPT (see above).

The GPT of plants mediates the import of Glc6-P into the heterotrophic plastids of sink tissues. GPT is expressed in nongreen tissues and serves to provide the Glc6-P, which is in turn channeled into the starch biosynthetic pathway or used to produce reduction equivalents via the oxidative pentose phosphate pathway (OPPP; Kammerer et al. 1998; Niewiadomski et al. 2005). In the red alga *G. sulphuraria*, a member of the pPTs was identified as a putative homologue of the plant GPT. Uptake experiments with reconstituted liposomes containing the recombinant protein showed that Glc6-P is not a relevant substrate, confirming previous studies using reconstituted total membranes (Weber et al. 2004; Linka et al. 2008). This poses the question as to how *G. sulphuraria* and red algae in general provide reduced carbon to the heterotrophic growing plastids. Interestingly, while in plants the activity of the fructose-1,6-bisphosphatase (FBPase) is redox regulated and inactivated by oxidation, the *G. sulphuraria* homologue is not (Reichert et al. 2003). *G. sulphuraria* can in this way supply its requirements of hexose phosphates during the night or heterotrophic growth conditions by import of triose phosphates mediated by the TPT (Linka et al. 2008).

The XPT displays a high sequence similarity to GPT, and the lack of introns in the *Arabidopsis* homologue suggests that it derived from the GPT by retrotranscription and genome insertion (Knappe et al. 2003). It exchanges TPs, 3-PGA, xylulose-5-P (Xyl5-P), and, to a lesser extent, erythrose-4-P (Ery4-P) and ribulose-5-P (Ru5-P). In *Arabidopsis*, the proposed function of XPT is to import Xyl5-P into the stroma where it is further integrated into the Calvin–Benson cycle and the OPPP. Another proposed function of the XPT is the replenishment of carbon skeleton intermediates withdrawn from the Calvin–Benson cycle and the OPPP and used for other biosynthetic pathways such as the synthesis of nucleotides and the shikimate pathway requiring Rib5-P and Ery4-P, respectively (Eicks et al. 2002). As for the GPT, XPT activity is not present in red algae (Linka et al. 2008).

PPT mediates the counter-exchange of cytosolic phospho*enol*pyruvate with P_i (Fischer et al. 1997). Chloroplasts lack a complete glycolytic pathway; therefore they depend on the import of PEP from the cytosol to fuel the shikimate pathway which starts with the reaction of PEP with Ery4-P coming from the pentose phosphate pathway (Borchert et al. 1993; Prabhakar et al. 2009). The shikimate pathway is a plastidic-localized route linking the metabolism of carbohydrates to the synthesis of aromatic compounds such as amino acids and diverse secondary metabolites (Herrmann and Weaver 1999). In red algae, the orthologue gene of *G. sulphuraria* displays similar substrate specificity and kinetic constants as its green counterpart, indicating that rhodoplasts also depend on PPT activity to drive the plastid-localized PEP-dependent reactions (Linka et al. 2008).

The Case of Glaucophytes

For the third lineage of Archaeplastida, the glaucophytes, knowledge on the carbon partitioning between the muroplast and the cytosol is rather limited. Transport assays with isolated muroplasts showed that phosphate uptake was inhibited by 3-PGA and DHAP, suggesting a mechanism of phosphate transport similar to that of higher plants. Moreover, phosphate uptake was inhibited by the hexose phosphate Glc6-P, similar to the transport mechanism acting in heterotrophic tissues such as in the guard cells of stomata and in the amyloplasts (Schlichting and Bothe 1993). Kinetic constants of Glc6-P transport were similar to those measured for triose phosphates, clearly showing that both metabolites are true substrates of the cyanelle phosphate translocator(s). The genetic identity of these translocators remained elusive for the last two decades, until the Cyanophora Genome Project threw light on the putative complement of the glaucophytes pPTs (http://dblab.rutgers.edu/cyanophora/; Price et al. 2012). Among a total of 27,921 proteins predicted to be coded by the *Cyanophora* genome, six genes were identified that shared a discrete homology to the NST family of transporters and could thus represent genuine candidates for the pPTs. If this were the case, the long sought phosphate translocator of the glaucophytes could provide additional support to the metabolic symbiosis hypothesis, which states that establishment of a favorable partnership between the symbiotic partners was put forward by the merging of the two metabolic entities and the photosynthate transporter would have played a crucial role in this process (Deschamps et al. 2008a). Surprisingly, phylogenetic analysis of the six candidates revealed that they do not localize in the branches of the tree which contain the pPTs of the red and green lineages (Price et al. 2012). It is therefore possible that glaucophytes possess transporters devoted to carbon export from the muroplast that are not related to the pPTs common to the other lineages. To this respect, a search for genes encoding putative plastidic carbon transporters retrieved two candidates that display similarity with the bacterial UhpC-type hexose phosphate transporters. Both genes encode a protein with an N-terminal extension containing a phenylalanine residue in a conserved position typical for *Cyanophora* plastid-targeted proteins (Steiner and Löffelhardt 2002; Steiner et al. 2005). The *Cyanophora* UhpC orthologues are related to the hexose phosphate translocators of Chlamydiae and *Legionella*, parasitic bacteria which are thought to have significantly contributed to the primary endosymbiosis (see below).

UhpC homologues are present in all three lineages originating from primary endosymbiosis, but they are restricted to the unicellular algae. It is therefore tempting to speculate that, due to the lack of true pPTs, glaucophytes represent a primitive stage during the evolution of the carbon allocation mechanism. Under this view, the recruitment of a hexose phosphate transporter must have preceded the insertion of the host-derived NST-type translocator. Remarkably, under this scenario the transfer of an ADP-glucose-dependent starch synthase to the cytosol

would have been dispensable since hexose phosphates can directly be channeled into the host's glycogen biosynthetic pathway. If this holds true, the appearance of the ADP-glucose translocator has to be dated after the split of the glaucophytes.

Other Transporters Involved in Carbon Partitioning

Besides the broadly distributed pPTs, Chloroplastida have also evolved a specific set of carbon transporters particularly suited to cope with the localization of starch in these organisms. As already mentioned, green algae and land plants accumulate their storage polysaccharide inside the chloroplast. At night, the starch reserve accumulated during the day is degraded and the breakdown products, namely, maltose and glucose, are exported to supply the cytosol with substrates for sucrose biosynthesis. The pathway of starch degradation starts with the release of linear glucans by the subsequent actions of the phosphorylating enzymes glucan water dikinase (GWD) and phosphoglucan water dikinase (PDW) and a debranching enzyme. Maltose is then released by β-amylases acting on linear glucans and, due to the inability of β-amylases to act on chains of less than four glucosyl residues, a smaller amount of maltotriose is produced which is further metabolized to glucose by the disproportionating enzyme DPE1 (Critchley et al. 2001; Smith et al. 2005). Maltose represents the predominant form by which photoassimilates leave the chloroplast during the dark or under increased sink demand and its export is facilitated by the MEX transporter identified by the maltose excess phenotype of plants deficient in its activity (Niittylä et al. 2004; Weise et al. 2004). The glucose transporter (pGlcT) also contributes to the export of the products of starch breakdown although it does not define the primary route for the export of starch degradation products (Weber et al. 2000). In addition, it has a not yet identified accessory function suggested by its expression in heterotrophic tissues devoid of starch (Butowt et al. 2003). Recent investigations using pglct-1/mex-1 double knockout *Arabidopsis* plants could confirm the importance of this transporter in plant growth and fertility (Hahn et al. 2011).

The MEX protein belongs to a small gene family consisting only of the maltose exporter, while pGlcT belongs to the major facilitator superfamily whose members are ubiquitously distributed among all the kingdoms. However, close relatives of this transporter are only found within the Chloroplastida. Their exclusive presence in the green algae and land plants reflects a late acquisition by the green lineage prompted by the redirection of the starch biosynthesis to the chloroplast with the consequent need to export its degradation products under heterotrophic conditions.

Nitrogen Metabolism

The main source of inorganic nitrogen available for assimilation is from elementary nitrogen (N_2) through the process of nitrogen fixation carried out by prokaryotic species converting gaseous nitrogen into ammonia (NH_4^+). Alternatively, ammonia

can be generated by assimilatory nitrate (NO_3^-) reduction by the concerted action of the cytosolic nitrate reductase (NR) and the plastidic nitrite reductase (NiR). Inorganic nitrogen is then incorporated into organic compounds via the plastidic glutamine synthetase/glutamate synthase cycle (GS/GOGAT). The net product of ammonium assimilation is one molecule of glutamate from one molecule of 2-oxoglutarate and ammonia each. Glutamate is the universal organic nitrogen donor delivering amino groups to the general metabolism. The GS/GOGAT system is also involved in the re-assimilation of the ammonia produced by the glycine decarboxylase in the mitochondrion during photorespiration (Renne et al. 2003; Linka and Weber 2005; Schneidereit et al. 2006). Ammonium is combined with glutamate by the GS in an ATP-dependent reaction, yielding glutamine which in turn reacts with 2-oxoglutarate resulting into two glutamate molecules. This reaction is catalyzed by the GOGAT and requires reduction equivalents. Incorporation of ammonia into glutamate requires a plastidic malate-coupled two-translocator system consisting of two dicarboxylate translocators (DiTs). DiT1 imports the acceptor molecule for ammonium assimilation, 2-oxoglutarate, in exchange with malate; DiT2 exports the end product of ammonia assimilation, glutamate, to the cytosol, importing back the malate (Weber et al. 1995; Weber and Flügge 2002; Renne et al. 2003). Phylogenetic analysis indicated that DiTs originated from a chlamydial ancestor by HGT (Tyra et al. 2007). In the Archaeplastida, DiTs are found only among green algae and land plants, but not in red algae and glaucophytes. In red algae, GS is encoded by the nuclear genome, whereas GOGAT is plastid encoded, indicating that the protein is localized in the plastid (Glöckner et al. 2000; Ohta et al. 2003). In addition, while land plants possess both cytosolic and plastidic isoforms of GS, red algal genomes code only for one cytosolic GS. This was explained by the strong affinity of the form IB RubisCO for CO_2 and low oxygenase activity in the Cyanidiales, a plausible reason for the dispensability of organelle GS due to low photorespiratory rates (Uemura et al. 1997; Terashita et al. 2006). Red algae likely possess an alternative and yet unknown transporter mediating the export of assimilated nitrogen from the plastid. In the case of glaucophyte algae, isolated muroplasts from *C. paradoxa* displayed a transport activity for glutamine and 2-oxoglutarate, whereas glutamate poorly penetrated into the plastid (Kloos et al. 1993). Ammonia formed by nitrite reduction inside the cyanelles is incorporated by GS into glutamine and then exported jointly with 2-oxoglutarate either by a single carrier or by separate ones (Kloos et al. 1993).

Taken together, the current information about the transporters involved in the shuttling of assimilated nitrogen across the plastid envelope membrane points to diverse mechanisms adopted by the three lineages of Archaeplastida, explainable either with the loss of the chlamydial dicarboxylate transporter from the red algae and the glaucophytes or with an independent acquisition of the latter by the green lineage after the split of chlorophyta and rhodophyta (Tyra et al. 2007).

Nitrogen Fixation

As mentioned before, the endosymbiont was likely an ancestor of modern-day group V cyanobacteria able to synthesize a starch-like polymer (Deschamps et al. 2008a). Moreover, these cyanobacteria are able to fix gaseous dinitrogen by means of the dinitrogenase enzyme. The enzyme is highly sensitive to oxygen and many filamentous cyanobacteria able to form colonies have evolved specialized structures, the heterocysts, devoted to nitrogen fixation where the nitrogenase is maintained in an oxygen-poor environment due to additional cell walls impermeable to oxygen and the lack of photosystem II. Unicellular non-colony-forming diazotrophic cyanobacteria evolved an adaptation mechanism which, instead of spatially separating the nitrogen fixation, separates it temporally. The genus *Cyanothece*, for example, has the ability to carry out aerobic nitrogen fixation by creating a microaerophilic intracellular environment at night, allowing oxygen-sensitive processes to take place. Carbohydrate reserves accumulate during the day, providing the energy necessary for the processes that require anoxic conditions of the cells (Bergman et al. 1997; Bandyopadhyay et al. 2011). Nitrogenase also produces hydrogen gas as part of its catalytic cycle. Hydrogen can leave the cell by diffusion, but nitrogen-fixing cyanobacteria are able to recycle the electrons by the enzyme uptake hydrogenase. This has a beneficial impact on the organism, since electrons are reintroduced into the electron transport chain thus contributing to the ATP production and, importantly, it provides ATP via the oxyhydrogen reaction where O_2 acts as final electron acceptor, thereby lowering the oxygen levels (Tamagnini et al. 2002; Bothe et al. 2010). Finally, the nitrogenase reaction, though occurring at ambient temperature and pressure (while the Haber–Bosch process requires much more extreme conditions), still is highly energy demanding: 16 molecules of ATP for every molecule of gaseous nitrogen fixed. Starting from this, possible benefits have been proposed for a diazotrophic cyanobacterium interacting with the host of the primary endosymbiotic event (Linka and Weber 2012). The close association with an heterotrophic eukaryote could have had beneficial effects on nitrogen fixation by providing a shielded, oxygen-poor environment favorable for the nitrogenase, by scavenging the molecular hydrogen, with the host cell acting as sink for the hydrogen derived by the nitrogenase, in accordance with the hydrogen hypothesis, and by eventually returning back the product of the hydrogen-based metabolism in form of ATP to sustain nitrogen fixation (Martin and Müller 1998; Linka and Weber 2012). However, it is still questionable whether the oxygen concentration could have reached levels low enough for the nitrogenase to work. The loss of carbon stores from the endosymbiont and the net export of carbohydrates implicated the loss of the O_2-consuming respiratory pathway which in the dark would have contributed to reach the required anoxia for the nitrogen fixation. It is thus likely that the ability to fix dinitrogen was lost from the endosymbiont simultaneously with the capacity to synthesize and store starch.

ATP Supply to the Cyanobiont by the Host

Supply of ATP requires active transport from the host cytosol to the cyanobacterium. Plastids possess an adenylate translocator which is unrelated to the mitochondrial ATP/ADP transporters and belongs to the major facilitator superfamily (Schünemann et al. 1993; Winkler and Neuhaus 1999). The ATP/ADP antiporter of plastids, or nucleoside triphosphate transporter (NTT), mediates the import of ATP into heterotrophic plastids to supply ATP-dependent reactions in nonphotosynthetic tissues. In heterotrophic plastids of sink tissues, such as amyloplasts, ATP import serves to drive ATP-dependent anabolic pathways like starch and amino acid biosynthesis and ammonia assimilation. Transgenic potato lines showed that tubers with decreased plastidic ATP/ADP transporter activities exhibited reduced starch contents, whereas overexpression lines accumulated increased amounts of starch (Tjaden et al. 1998). On the other hand, in photosynthetic plastids NTT serves as ATP importer at night or during increasing energy demand (Reinhold et al. 2007). The assembly of magnesium chelatase, an enzyme involved in chlorophyll biosynthesis, depends on ATP import (Kobayashi et al. 2008). *Arabidopsis* plants with mutations in the genes encoding two plastidic ATP/ADP transporters display necrotic lesions caused by photooxidation due to an accumulation of high levels of phototoxic protoporphyrin IX, responsible for ROS production and photooxidative damage. The necrotic phenotype was rescued by conditions of long days and low light intensity or short days and high light intensity (Reinhold et al. 2007). High levels of stromal ATP are required for detoxification of the cytotoxic intermediate. Deschamps et al. (2008c) proposed that, alternatively, ATP could be produced inside the stroma by starch degradation. Starch phosphorylase produces hexose phosphates, which are in turn further metabolized to generate the ATP pool necessary to alleviate the protoporphyrin IX-induced oxidative stress. This is proposed to be the reason for restoring the synthesis of starch inside the chloroplast as protective mechanism for a safe evolution of the LHCs (Deschamps et al. 2008c).

In *Cyanophora*, a genomic clone has been identified with a distinct sequence similarity to the NTTs, even if uptake experiments with intact muroplasts could not show ATP transport activity (Schlichting et al. 1990; Linka et al. 2003). Of the two major ATP-consuming plastidic pathways (starch and fatty acid biosynthesis) only fatty acid biosynthesis is located to the plastid in *Cyanophora*. ATP requirements in the muroplast could thus possibly be sustained by substrate-level phosphorylation in the absence of photosynthetic electron transport.

In red algae, homologues of the NTTs are found in the genomes of *Cyanidioschyzon merolae* and *G. sulphuraria* and, at least for *Galdieria*, transport experiments could prove the ATP/ADP exchange activity (Linka et al. 2003; Tyra et al. 2007). As for glaucophytes, red algae store their carbon in the cytosol and fatty acids biosynthesis occurs in the plastid, as indicated by the presence of the subunits of the plastidic acetyl-CoA carboxylase, requiring ATP (Weber et al. 2004).

The fact that the NTTs are encoded by the genome of all three members of Archaeplastida points to a very ancient origin of the transporter, before the split of the lineages. Besides in plants' plastids, NTTs are only found in bacteria. Here they function to take up ATP from the eukaryotic cell cytoplasm into the bacterium in exchange for ADP and enable energy parasitism in *Chlamydia* and *Rickettsia* (Neuhaus et al. 1997; Winkler and Neuhaus 1999).

The early appearance of the NTTs is consistent with a role in the initial stage of the endosymbiotic association. As denoted before, ATP supply to the cyanobacterium could have favored a more efficient nitrogen fixation. Targeting of an adenylate translocator to the endosymbiont's membrane would have certainly contributed significantly to the successful integration of the cyanobiont and in this way prompted the interdependence of the two partners.

Evolutionary Origin of the Plastidic Translocators

The establishment of a metabolic connection was of crucial importance for the success of endosymbiosis. As seen for the members of the pPT family, the setting up of a carbon flux was a process driven by the host by recruiting a translocator from its endomembrane system and redirecting it to the endosymbiont. But what about the other proteins responsible for the exchange of metabolites across the envelope? Tyra et al. (2007) addressed the question as to how the plastid solute transport system was established by taking a set of 137 *Arabidopsis* solute transporters and performing a detailed phylogenetic analysis of a subset of 83 proteins conserved and broadly distributed among Plantae. The analysis revealed that the integration of the two metabolic units was driven by the host, as 58 % of the transporters scrutinized were found to be of host origin. Besides the above-mentioned nucleotide sugar/triose phosphate translocator gene family, all the carbohydrate transporters were assigned as being of host origin. This indicates that the host adopted its eukaryotic transport system to draw off photosynthates from the captured cyanobacterium. In addition, where a cyanobacterial transporter was present, this was replaced by a host version, presumably to favor a rapid acquisition of the captured endosymbiont (Tyra et al. 2007). According to this scenario, the cyanobacteria-derived transporters that were retained in the envelope (representing a mere 12 % of the total) are those for which a host version was lacking. A third contributor with 8 % of the transporters analyzed is Chlamydiae. These include the above-mentioned NTTs, the DiTs, a phosphate transporter, and a heavy metal ATPase.

Intracellular Parasites as Drivers for the Metabolic Integration

It is surprising that intracellular bacteria that do not occur in modern plant cells contributed to such an extent to the transporter complement during plastid evolution. An analysis aimed at detecting bacterial proteins more similar in primary sequence to eukaryotic proteins over other bacterial or archaeal proteins (and vice versa) showed that 65 % of bacterial proteins identified with the highest similarity to a eukaryotic protein involved *Chlamydia*, *Chlamydophila*, *Synechocystis*, and *Rickettsia*, although these organisms only accounted for 14 % of the genes analyzed. The proteins identified from *Rickettsia* were found to be mostly belonging to the "energy production and conversion" functional category and the *Synechocystis* and Chlamydiaceae proteins were found to be more similar to plant proteins (Brinkman et al. 2002). Although for *Rickettsia* (an α-proteobacterium) and *Synechocystis* this was not a surprise, due to the ancestral relationship of these with the mitochondrion and the chloroplast, respectively, the presence of plant-like genes in Chlamydiaceae was intriguing, also because for the majority of these genes the plant counterpart contained a plastid-targeting signal (Brinkman et al. 2002).

Chlamydiae are obligate intracellular bacteria which can either behave as pathogens, such as in humans and animals, or as endosymbionts in environmental amoebae, which are believed to represent reservoirs (or "Trojan horses") for obligate intracellular bacteria (Horn and Wagner 2004). Even if the number of identified environmental Chlamydiae is constantly increasing, to date no chlamydial species have been reported in plastid-bearing eukaryotes (Horn and Wagner 2001).

The occurrence of genes of chlamydial origin in photosynthetic eukaryotes and non-photosynthetic plastid-containing lineages suggests an intimate association between Chlamydiae and the plant ancestor. In order to test this, the phylogeny of Chlamydiae and the red alga *C. merolae* was conducted to search for genes that are evolutionarily related (Huang and Gogarten 2007). The analysis retrieved a list of 21 Chlamydiae-related genes, most of which are also present in green plants, indicating that the bacterial contribution spans throughout the primary photosynthetic eukaryotes. The same study also provided an answer to the question about the direction of the transfer: previous assumptions of a plant-to-Chlamydia transfer are no longer supported since the cyanobacterial homologues form a well-supported branch distinct from the chlamydial homologues (Huang and Gogarten 2007). Deeper phylogenomic analyses of more plant genomes and comparison with a large bacterial dataset further extended the list of proteins of chlamydial origin (Becker et al. 2008; Moustafa et al. 2008). The relative high number of sequences of chlamydial origin found in plant genomes may be explained by a long-term endosymbiotic relationship involving EGT rather than multiple HGT. This is corroborated by the observation that the majority of the Plantae genes identified by Moustafa et al. (2008), namely 30 out of 55, show a specific relationship with the environmental Chlamydia UWE25 (*Candidatus Protochlamydia amoebophila*). A way of gene transfer is also provided by a type IV secretion system (TFSS) present

in the genome of UWE25 and absent in pathogenic bacteria, which could play a role in conjugative DNA transfer (Horn et al. 2004). It is therefore plausible that an ancient association between Chlamydiae and the ancestor of primary photosynthetic eukaryotes provided the latter with some key genes and in particular transporters played a crucial role. Chlamydiae were likely already present in the cytosol of the eukaryotic host before the entry of the cyanobacterium and adapted to drain ATP from the host by the ATP/ADP translocase. Insertion of this translocator into the cyanobacterial membrane represented probably a major advance in the connection of the two partners, enabling the equilibration of their ATP pool and, as mentioned previously, could have offered a specific advantage for the endosymbiont in the energy-demanding process of nitrogen fixation. Similarly, the DiTs of chlamydial origin are also involved in assimilating nitrogen. Most intriguing is the presence of a hexose phosphate translocator related to the UhpC transporter from Chlamydiae, which, if confirmed to reside in the envelope membrane and to catalyze the export of hexose phosphates, could date back the evolution of a carbon export mechanism before the appearance of the host-derived NST-type translocator.

Conclusions

The process of plastid acquisition through primary endosymbiosis was a unique event which occurred only once during evolution and gave rise to the ancestor of all photoautotrophic eukaryotes. These were responsible for the conquering of the land and for an enormous biodiversity. To date, this unique, ancient event encounters only one other much more recent parallel in the case of the "plastid in the making" of the thecate amoeba *Paulinella chromatophora* (Marin et al. 2005; Nowack et al. 2008).

The transport of metabolites across the endosymbiont's membrane was decisive for the establishment of a successful endosymbiosis in the algal ancestor. The picture emerging from genomic, phylogenomic, and biochemical data greatly contributes to deciphering the events that may have occurred at the very early stages of metabolic integration. These point to a third, until recently unappreciated, contribution from chlamydial endoparasites, which very likely inhabited the eukaryotic host before the enslavement of the cyanobacterium (Fig. 2). These bacteria contributed through EGT to a relatively small but essential set of genes. The most important ones specify transporters that are now involved in the major plastidic metabolic pathways. Later on, endoparasitic Chlamydiae as well as the ability to fix nitrogen were lost from phototrophic eukaryotes.

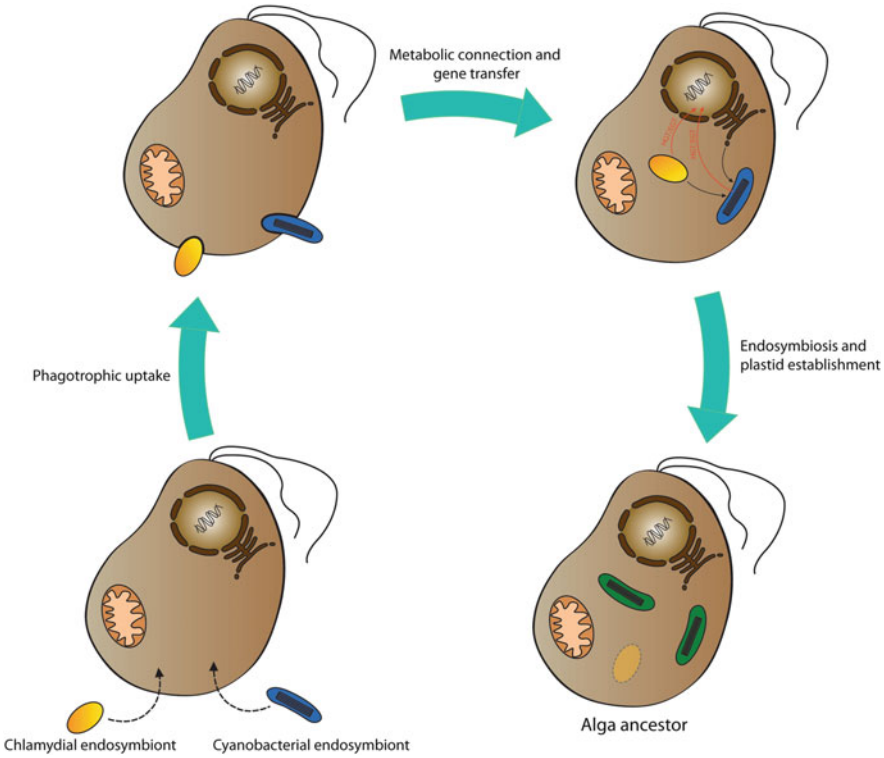


Fig. 2 Hypothetical scenario for plastid origin and establishment. The mitochondriate eukaryote was a phagotroph feeding on cyanobacteria. Chlamydiae were likely already present in the host's cytosol as intracellular parasites. The chlamydial endosymbiont already held the transport system necessary to parasitize the host cell. The tight association of the three organisms favored the gene transfer from both prokaryotes to the host nucleus. Communication of the latter with the cyanobacterial captive was facilitated by the transporters acquired from the chlamydial endosymbiont. The acquisition of transporters of host origin contributed to establish a metabolic connection followed by massive gene transfer to the nucleus along with the evolution of a protein import apparatus. As a consequence, the cyanobacterium gradually turned into an organelle, while the chlamydial parasite would have been lost (the selective advantage of autotrophy could only be conferred by the cyanobacterial endosymbiont)

References

Adl SM, Simpson AGB, Farmer MA, Andersen RA, Anderson OR, Barta JR, Bowser SS, Brugerolle G, Fensome RA, Fredericq S, James TY, Karpov S, Kugrens P, Krug J, Lane CE, Lewis LA, Lodge J, Lynn DH, Mann DG, Mccourt RM, Mendoza L, Moestrup O, Mozley-Standridge SE, Nerad TA, Shearer CA, Smirnov AV, Spiegel FW, Taylor MFJR (2005) The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. J Eukaryot Microbiol 52:399–451

Allen JF (2005) A redox switch hypothesis for the origin of two light reactions in photosynthesis. FEBS Lett 579:963–968

Allen JF, Martin W (2007) Evolutionary biology - out of thin air. Nature 445:610–612

Ball SG, Morell MK (2003) From bacterial glycogen to starch: understanding the biogenesis of the plant starch granule. Annu Rev Plant Biol 54:207–233

Ball S, Colleoni C, Cenci U, Raj JN, Tirtiaux C (2011) The evolution of glycogen and starch metabolism in eukaryotes gives molecular clues to understand the establishment of plastid endosymbiosis. J Exp Bot 62:1775–1801

Bandyopadhyay A, Elvitigala T, Welsh E, Stockel J, Liberton M, Min HT, Sherman LA, Pakrasi HB (2011) Novel metabolic attributes of the genus cyanothece. Comprising a group of unicellular nitrogen-fixing cyanobacteria. MBio 4:2(5)

Becker B, Hoef-Emden K, Melkonian M (2008) Chlamydial genes shed light on the evolution of photoautotrophic eukaryotes. BMC Evol Biol 8:203

Bergman B, Gallon JR, Rai AN, Stal LJ (1997) N_2 fixation by non-heterocystous cyanobacteria. FEMS Microbiol Rev 19:139–185

Bhattacharya D, Yoon HS, Hackett JD (2004) Photosynthetic eukaryotes unite: endosymbiosis connects the dots. Bioessays 26:50–60

Borchert S, Harborth J, Schünemann D, Hoferichter P, Heldt HW (1993) Studies of the enzymic capacities and transport properties of pea root plastids. Plant Physiol 101:303–312

Bothe H, Schmitz O, Yates MG, Newton WE (2010) Nitrogen fixation and hydrogen metabolism in cyanobacteria. Microbiol Mol Biol Rev 74:529–551

Brinkman FSL, Blanchard JL, Cherkasov A, Av-Gay Y, Brunham RC, Fernandez RC, Finlay BB, Otto SP, Ouellette BFF, Keeling PJ, Rose AM, Hancock REW, Jones SJM (2002) Evidence that plant-like genes in Chlamydia species reflect an ancestral relationship between Chlamydiaceae, cyanobacteria, and the chloroplast. Genome Res 12:1159–1167

Bryant DA, Frigaard NU (2006) Prokaryotic photosynthesis and phototrophy illuminated. Trends Microbiol 14:488–496

Busch A, Nield J, Hippler M (2010) The composition and structure of photosystem I-associated antenna from *Cyanidioschyzon merolae*. Plant J 62:886–897

Butowt R, Granot D, Rodriguez-Garcia MI (2003) A putative plastidic glucose translocator is expressed in heterotrophic tissues that do not contain starch, during olive (Olea europea L) fruit ripening. Plant Cell Physiol 44:1152–1161

Cavalier-Smith T (1982) The origins of plastids. Biol J Linn Soc Lond 17:289–306

Cavalier-Smith T (1999) Principles of protein and lipid targeting in secondary symbiogenesis: euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree. J Eukaryot Microbiol 46:347–366

Cavalier-Smith T (2000) Membrane heredity and early chloroplast evolution. Trends Plant Sci 5:174–182

Cavalier-Smith T (2003) Genomic reduction and evolution of novel genetic membranes and protein-targeting machinery in eukaryote-eukaryote chimaeras (meta-algae). Philos Trans R Soc Lond B Biol Sci 358:109–133

Colleoni C, Linka M, Deschamps P, Handford MG, Dupree P, Weber APM, Ball SG (2010) Phylogenetic and biochemical evidence supports the recruitment of an ADP-glucose translocator for the export of photosynthate during plastid endosymbiosis. Mol Biol Evol 27:2691–2701

Critchley JH, Zeeman SC, Takaha T, Smith AM, Smith SM (2001) A critical role for disproportionating enzyme in starch breakdown is revealed by a knock-out mutation in *Arabidopsis*. Plant J 26:89–100

Deschamps P, Colleoni C, Nakamura Y, Suzuki E, Putaux JL, Buleon A, Haebel S, Ritte G, Steup M, Falcon LI, Moreira D, Löffelhardt W, Raj JN, Plancke C, D'hulst C, Dauvillee D, Ball S (2008a) Metabolic symbiosis and the birth of the plant kingdom. Mol Biol Evol 25:536–548

Deschamps P, Haferkamp I, D'hulst C, Neuhaus HE, Ball SG (2008b) The relocation of starch metabolism to chloroplasts: when, why and how. Trends Plant Sci 13:574–582

Deschamps P, Moreau H, Worden AZ, Dauvillee D, Ball SG (2008c) Early gene duplication within chloroplastida and its correspondence with relocation of starch metabolism to chloroplasts. Genetics 178:2373–2387

Duy D, Soll J, Philippar K (2007) Solute channels of the outer membrane: from bacteria to chloroplasts. Biol Chem 388:879–889

Eicks M, Maurino V, Knappe S, Flügge UI, Fischer K (2002) The plastidic pentose phosphate translocator represents a link between the cytosolic and the plastidic pentose phosphate pathways in plants. Plant Physiol 128:512–522

Facchinelli F, Weber APM (2011) The metabolite transporters of the plastid envelope: an update. Front Plant Sci 2:50

Falkowski PG, Godfrey LV (2008) Electrons, life and the evolution of Earth's oxygen cycle. Philos Trans R Soc Lond B Biol Sci 363:2705–2716

Fischer K, Kammerer B, Gutensohn M, Arbinger B, Weber A, Häusler RE, Flügge UI (1997) A new class of plastidic phosphate translocators: a putative link between primary and secondary metabolism by the phosphoenolpyruvate/phosphate antiporter. Plant Cell 9:453–462

Fliege R, Flügge UI, Werdan K, Heldt HW (1978) Specific transport of inorganic-phosphate, 3-phosphoglycerate and triosephosphates across inner membrane of envelope in spinach-chloroplasts. Biochim Biophys Acta 502:232–247

Flügge UI (1998) Metabolite transporters in plastids. Curr Opin Plant Biol 1:201–206

Flügge UI (2000) Transport in and out of plastids: does the outer envelope membrane control the flow? Trends Plant Sci 5:135–137

Flügge UI, Heldt HW (1984) The phosphate-triose phosphate-phosphoglycerate translocator of the chloroplast. Trends Biochem Sci 9:530–533

Flügge UI, Häusler RE, Ludewig F, Fischer K (2003) Functional genomics of phosphate antiport systems of plastids. Physiol Plant 118:475–482

Glöckner G, Rosenthal A, Valentin K (2000) The structure and gene repertoire of an ancient red algal plastid genome. J Mol Evol 51:382–390

Gould SB, Waller RF, Mcfadden GI (2008) Plastid evolution. Annu Rev Plant Biol 59:491–517

Gross J, Bhattacharya D (2009) Revaluating the evolution of the Toc and Tic protein translocons. Trends Plant Sci 14:13–20

Hahn TR, Cho MH, Lim H, Shin DH, Jeon JS, Bhoo SH, Park YI (2011) Role of the plastidic glucose translocator in the export of starch degradation products from the chloroplasts in *Arabidopsis thaliana*. New Phytol 190:101–112

Herrmann KM, Weaver LM (1999) The shikimate pathway. Annu Rev Plant Biol 50:473–503

Hohmann-Marriott MF, Blankenship RE (2011) Evolution of photosynthesis. Annu Rev Plant Biol 62:515–548

Horn M, Wagner M (2001) Evidence for additional genus-level diversity of Chlamydiales in the environment. FEMS Microbiol Lett 204:71–74

Horn M, Wagner M (2004) Bacterial endosymbionts of free-living amoebae. J Eukaryot Microbiol 51:509–514

Horn M, Collingro A, Schmitz-Esser S, Beier CL, Purkhold U, Fartmann B, Brandt P, Nyakatura GJ, Droege M, Frishman D, Rattei T, Mewes HW, Wagner M (2004) Illuminating the evolutionary history of chlamydiae. Science 304:728–730

Huang JL, Gogarten JP (2007) Did an ancient chlamydial endosymbiosis facilitate the establishment of primary plastids? Genome Biol 8:R99

Jain R, Rivera MC, Lake JA (1999) Horizontal gene transfer among genomes: the complexity hypothesis. Proc Natl Acad Sci U S A 96:3801–3806

Kammerer B, Fischer K, Hilpert B, Schubert S, Gutensohn M, Weber A, Flügge UI (1998) Molecular characterization of a carbon transporter in plastids from heterotrophic tissues: the glucose 6-phosphate phosphate antiporter. Plant Cell 10:105–117

Keeling PJ (2010) The endosymbiotic origin, diversification and fate of plastids. Philos Trans R Soc Lond B Biol Sci 365:729–748

Kloos K, Schlichting R, Zimmer W, Bothe H (1993) Glutamine and glutamate transport in Cyanophora-paradoxa. Bot Acta 106:435–440

Knappe S, Flügge UI, Fischer K (2003) Analysis of the plastidic phosphate translocator gene family in *arabidopsis* and identification of new phosphate translocator-homologous transporters, classified by their putative substrate-binding site. Plant Physiol 131:1178–1190

Kobayashi K, Mochizuki N, Yoshimura N, Motohashi K, Hisabori T, Masuda T (2008) Functional analysis of *Arabidopsis thaliana* isoforms of the Mg-chelatase CHLI subunit. Photochem Photobiol Sci 7:1188–1195

Lane N (2009) Life ascending: the ten great inventions of evolution. Profile Books, London

Linka M, Weber APM (2005) Shuffling ammonia between mitochondria and plastids during photorespiration. Trends Plant Sci 10:461–465

Linka N, Weber APM (2010) Intracellular metabolite transporters in plants. Mol Plant 3:21–53

Linka M, Weber APM (2012) Evolutionary integration of chloroplast metabolism with the metabolic networks of the cells. In: Burnap R, Vermaas W (eds) Functional genomics and evolution of photosynthetic systems. Springer, Netherlands, pp 199–224

Linka N, Hurka H, Lang BF, Burger G, Winkler HH, Stamme C, Urbany C, Seil I, Kusch J, Neuhaus HE (2003) Phylogenetic relationships of non-mitochondrial nucleotide transport proteins in bacteria and eukaryotes. Gene 306:27–35

Linka M, Jamai A, Weber APM (2008) Functional characterization of the plastidic phosphate translocator gene family from the thermo-acidophilic Red alga *Galdieria sulphuraria* reveals specific adaptations of primary carbon partitioning in green plants and Red algae. Plant Physiol 148:1487–1496

Löffelhardt W, Bohnert H, Bryant D, Hagemann PDR (1997) The cyanelles of Cyanophora paradoxa. CRC Crit Rev Plant Sci 16:393–413

Marin B, Nowack EC, Melkonian M (2005) A plastid in the making: evidence for a second primary endosymbiosis. Protist 156:425–432

Martin W, Müller M (1998) The hydrogen hypothesis for the first eukaryote. Nature 392:37–41

Martin W, Stoebe B, Goremykin V, Hansmann S, Hasegawa M, Kowallik KV (1998) Gene transfer to the nucleus and the evolution of chloroplasts. Nature 393:162–165

Mereschkowsky C (1905) Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. Biol Centralbl 25:593–604

Moreira D, Le Guyader H, Philippe H (2000) The origin of red algae and the evolution of chloroplasts. Nature 405:69–72

Moustafa A, Reyes-Prieto A, Bhattacharya D (2008) Chlamydiae Has contributed at least 55 genes to plantae with predominantly plastid functions. PLoS One 3(5):e2205

Neuhaus HE, Thom E, Mohlmann T, Steup M, Kampfenkel K (1997) Characterization of a novel eukaryotic ATP/ADP translocator located in the plastid envelope of *Arabidopsis thaliana* L. Plant J 11:73–82

Niewiadomski P, Knappe S, Geimer S, Fischer K, Schulz B, Unte US, Rosso MG, Ache P, Flügge UI, Schneider A (2005) The *Arabidopsis* plastidic glucose 6-phosphate/phosphate translocator GPT1 is essential for pollen maturation and embryo sac development. Plant Cell 17:760–775

Niittylä T, Messerli G, Trevisan M, Chen J, Smith AM, Zeeman SC (2004) A previously unknown maltose transporter essential for starch degradation in leaves. Science 303:87–89

Nowack EC, Melkonian M, Glöckner G (2008) Chromatophore genome sequence of Paulinella sheds light on acquisition of photosynthesis by eukaryotes. Curr Biol 18:410–418

Ohta N, Matsuzaki M, Misumi O, Miyagishima S, Nozaki H, Tanaka K, Shini T (2003) Complete sequence and analysis of the plastid genome of the unicellular red alga *Cyanidioschyzon merolae*. DNA Res 10:67–77

Olson JM (2006) Photosynthesis in the Archean Era. Photosynth Res 88:109–117

Olson JM, Blankenship RE (2004) Thinking about the evolution of photosynthesis. Photosynth Res 80:373–386

Plancke C, Colleoni C, Deschamps P, Dauvillee D, Nakamura Y, Haebel S, Ritte G, Steup M, Buleon A, Putaux JL, Dupeyre D, D'hulst C, Ral JP, Löffelhardt W, Ball SG (2008) Pathway of

cytosolic starch synthesis in the model glaucophyte Cyanophora paradoxa. Eukaryot Cell 7:247–257

Prabhakar V, Löttgert T, Gigolashvili T, Bell K, Flügge UI, Häusler RE (2009) Molecular and functional characterization of the plastid-localized phosphoenolpyruvate enolase (ENO1) from *Arabidopsis thaliana*. FEBS Lett 583:983–991

Price DC, Chan CX, Yoon HS, Yang EC, Qiu H, Weber APM, Schwacke R, Gross J, Blouin NA, Lane C, Reyes-Prieto A, Durnford DG, Neilson JAD, Lang BF, Burger G, Steiner JM, Löffelhardt W, Meuser JE, Posewitz MC, Ball S, Arias MC, Henrissat B, Coutinho BM, Rensing SA, Symeonidi A, Doddapaneni H, Green BR, Rajah VD, Boore J, Bhattacharya D (2012) *Cyanophora paradoxa* genome elucidates origin of photosynthesis in algae and plants. Science 335:843–847

Raven JA (2009) Functional evolution of photochemical energy transformations in oxygen-producing organisms. Funct Plant Biol 36:505–515

Reichert A, Dennes A, Vetter S, Scheibe R (2003) Chloroplast fructose 1,6-bisphosphatase with changed redox modulation: comparison of the *Galdieria* enzyme with cysteine mutants from spinach. Biochim Biophys Acta 1645:212–217

Reinhold T, Alawady A, Grimm B, Beran KC, Jahns P, Conrath U, Bauer J, Reiser J, Melzer M, Jeblick W (2007) Limitation of nocturnal import of ATP into *Arabidopsis* chloroplasts leads to photooxidative damage. Plant J 50:293–304

Renne P, Dressen U, Hebbeker U, Hille D, Flügge UI, Westhoff P, Weber APM (2003) The *Arabidopsis* mutant dct is deficient in the plastidic glutamate/malate translocator DiT2. Plant J 35:316–331

Rodríguez-Ezpeleta N, Brinkmann H, Burey SC, Roure B, Burger G, Löffelhardt W, Bohnert HJ, Philippe H, Lang BF (2005) Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes. Curr Biol 15:1325–1330

Russell M (2006) First life. Am Sci 94:32–39

Sadekar S, Raymond J, Blankenship RE (2006) Conservation of distantly related membrane proteins: photosynthetic reaction centers share a common structural core. Mol Biol Evol 23:2001–2007

Sagan L (1967) On the origin of mitosing cells. J Theor Biol 14:225–274

Sato M, Mogi Y, Nishikawa T, Miyamura S, Nagumo T, Kawano S (2009) The dynamic surface of dividing cyanelles and ultrastructure of the region directly below the surface in *Cyanophora paradoxa*. Planta 229:781–791

Schimper AFW (1883) Über die Entwicklung der Chlorophyllkörner und Farbkörper. Bot Zeitung 41:105–114, 121–131, 137–146, 153–162

Schleiff E, Eichacker LA, Eckart K, Becker T, Mirus O, Stahl T, Soll J (2003) Prediction of the plant beta-barrel proteome: a case study of the chloroplast outer envelope. Protein Sci 12:748–759

Schlichting R, Bothe H (1993) The cyanelles (organelles of a Low evolutionary scale) possess a phosphate-translocator and a glucose-carrier in *Cyanophora paradoxa*. Bot Acta 106:428–434

Schlichting R, Zimmer W, Bothe H (1990) Exchange of metabolites in *Cyanophora paradoxa* and its cyanelles. Bot Acta 103:392–398

Schlichting R, Kloos K, Zimmer W, Bothe H (1994) The metabolite transfer in *Cyanophora paradoxa* and its cyanelles. Endocytobiosis Cell Res 10:81–85

Schneegurt MA, Sherman DM, Nayar S, Sherman LA (1994) Oscillating behavior of carbohydrate granule formation and dinitrogen fixation in the cyanobacterium *Cyanothece* Sp strain ATCC-51142. J Bacteriol 176:1586–1597

Schneidereit J, Häusler RE, Fiene G, Kaiser WM, Weber APM (2006) Antisense repression reveals a crucial role of the plastidic 2-oxoglutarate/malate translocator DiT1 at the interface between carbon and nitrogen metabolism. Plant J 45:206–224

Schünemann D, Borchert S, Flügge UI, Heldt HW (1993) ADP/ATP translocator from pea root plastids - comparison with translocators from spinach-chloroplasts and pea leaf mitochondria. Plant Physiol 103:131–137

Shi T, Falkowski PG (2008) Genome evolution in cyanobacteria: the stable core and the variable shell. Proc Natl Acad Sci U S A 105:2510–2515

Smith AM, Zeeman SC, Smith SM (2005) Starch degradation. Annu Rev Plant Biol 56:73–98

Soll J, Bolter B, Wagner R, Hinnah SC (2000) Response: the chloroplast outer envelope: a molecular sieve? Trends Plant Sci 5:137–138

Steiner JM, Löffelhardt W (2002) Protein import into cyanelles. Trends Plant Sci 7:72–77

Steiner JM, Yusa F, Pompe JA, Löffelhardt W (2005) Homologous protein import machineries in chloroplasts and cyanelles. Plant J 44:646–652

Tamagnini P, Axelsson R, Lindberg P, Oxelfelt F, Wunschiers R, Lindblad P (2002) Hydrogenases and hydrogen metabolism of cyanobacteria. Microbiol Mol Biol Rev 66:1–20

Terashita M, Maruyama S, Tanaka K (2006) Cytoplasmic localization of the single glutamine synthetase in a unicellular red alga, *Cyanidioschyzon merolae* 10D. Biosci Biotechnol Biochem 70:2313–2315

Timmis JN, Ayliffe MA, Huang CY, Martin W (2004) Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nat Rev Genet 5:123–135

Tjaden J, Mohlmann T, Kampfenkel K, Henrichs G, Neuhaus HE (1998) Altered plastidic ATP/ADP-transporter activity influences potato (*Solanum tuberosum* L) tuber morphology, yield and composition of tuber starch. Plant J 16:531–540

Tyra HM, Linka M, Weber APM, Bhattacharya D (2007) Host origin of plastid solute transporters in the first photosynthetic eukaryotes. Genome Biol 8:R212

Uemura K, Anwaruzzaman Miyachi S, Yokota A (1997) Ribulose-1,5-bisphosphate carboxylase/oxygenase from thermophilic red algae with a strong specificity for CO2 fixation. Biochem Biophys Res Commun 233:568–571

Umena Y, Kawakami K, Shen JR, Kamiya N (2011) Crystal structure of oxygen-evolving photosystem II at a resolution of 19Å. Nature 473:55–60

Vanselow C, Weber APM, Krause K, Fromme P (2009) Genetic analysis of the photosystem I subunits from the red alga, *Galdieria sulphuraria*. Biochim Biophys Acta 1787:46–59

Viola R, Nyvall P, Pedersen M (2001) The unique features of starch metabolism in red algae. Proc R Soc Lond B Biol Sci 268:1417–1422

Weber APM, Fischer K (2007) Making the connections - the crucial role of metabolite transporters at the interface between chloroplast and cytosol. FEBS Lett 581:2215–2222

Weber A, Flügge UI (2002) Interaction of cytosolic and plastidic nitrogen metabolism in plants. J Exp Bot 53:865–874

Weber APM, Linka N (2011) Connecting the plastid: transporters of the plastid envelope and their role in linking plastidial with cytosolic metabolism. Annu Rev Plant Biol 62:53–77

Weber A, Menzlaff E, Arbinger B, Gutensohn M, Eckerskorn C, Flügge UI (1995) The 2-oxoglutarate malate translocator of chloroplast envelope membranes - molecular-cloning of a transporter containing a 12-helix motif and expression of the functional protein in yeast cells. Biochemistry 34:2621–2627

Weber A, Servaites JC, Geiger DR, Kofler H, Hille D, Groner F, Hebbeker U, Flügge UI (2000) Identification, purification, and molecular cloning of a putative plastidic glucose translocator. Plant Cell 12:787–801

Weber APM, Oesterhelt C, Gross W, Brautigam A, Imboden LA, Krassovskaya I, Linka N, Truchina J, Schneidereit J, Voll H, Voll LM, Zimmermann M, Jamai A, Riekhof WR, Yu B, Garavito RM, Benning C (2004) EST-analysis of the thermo-acidophilic red microalga *Galdieria sulphuraria* reveals potential for lipid A biosynthesis and unveils the pathway of carbon export from rhodoplasts. Plant Mol Biol 55:17–32

Weber APM, Schwacke R, Flügge UI (2005) Solute transporters of the plastid envelope membrane. Annu Rev Plant Biol 56:133–164

Weber APM, Linka M, Bhattacharya D (2006) Single, ancient origin of a plastid metabolite translocator family in Plantae from an endomembrane-derived ancestor. Eukaryot Cell 5:609–612

Weise SE, Weber APM, Sharkey TD (2004) Maltose is the major form of carbon exported from the chloroplast at night. Planta 218:474–482

Winkler HH, Neuhaus HE (1999) Non-mitochondrial ATP transport. Trends Biochem Sci 24:64–68

Wolfe GR, Cunningham FX, Durnford D, Green BR, Gantt E (1994) Evidence for a common origin of chloroplasts with light-harvesting complexes of different pigmentation. Nature 367:566–568

Yoon HS, Hackett JD, Ciniglia C, Pinto G, Bhattacharya D (2004) A molecular timeline for the origin of photosynthetic eukaryotes. Mol Biol Evol 21:809–818

Evolution of the Protein Translocons of the Chloroplast Envelope

Maik S. Sommer and Enrico Schleiff

Abstract It is widely accepted that chloroplasts originated through the endosymbiotic uptake of an ancestral cyanobacterium by a mitochondria-containing heterotrophic host cell. In the course of evolution these once autonomous bacteria became increasingly integrated into their host's cellular environment. The permanent transfer of large portions of genetic information from the endosymbiont to the host's genome was essential for the establishment of a truly symbiotic relationship. However, as a consequence, the import of now cytoplasmically synthesized proteins into these new organelles became essential and specific translocons for the import of these proteins had to evolve. Interestingly, evidence suggests that chloroplasts take advantage of already existing prokaryotic proteins as the import machinery evolved. Here, we discuss the phylogenetic relationships of known translocon components and try to reconstruct important steps in the evolution of the import machinery.

Contents

The Complexity of Protein Translocation into Chloroplasts 82
The General Protein Translocon in Land Plants 83
The Various Translocation Routes Across Prokaryotic Membranes 85
The Evolution of the Targeting and Translocation Mechanisms into Plastids 88
The Possible Origin of the Targeting Mechanism 88
The Targeting Signals and the Targeting Complexes 89
The Evolution of the Outer Envelope Translocon 94

M.S. Sommer • E. Schleiff (✉)
Cluster of Excellence Macromolecular Complexes, Centre of Membrane Proteomics and Institute for Molecular Biosciences, Department of Biosciences, Molecular Cell Biology of Plants, Goethe University Frankfurt am Main, Max-von-Laue Str. 9, 60438 Frankfurt, Germany
e-mail: schleiff@bio.uni-frankfurt.de

W. Löffelhardt (ed.), *Endosymbiosis*, DOI 10.1007/978-3-7091-1303-5_5,
© Springer-Verlag Wien 2014

The Evolution of the Inner Envelope Translocon 99
The Evolution of the Outer Envelope Protein Insertion Machinery 100
Conclusion 101
References 102

The Complexity of Protein Translocation into Chloroplasts

The study of protein translocation into plastids of land plants has led to a general understanding of some of the factors involved in the translocation process. Initially it was thought that chloroplasts exclusively utilize the so-called "general translocation pathway." Proteins that need to be imported into the chloroplast, termed *PRECURSOR PROTEINS* or *PREPROTEINS*, typically contain an N-terminal extension (*TRANSIT PEPTIDE*) which serves as a signal and is cleaved upon arrival in the stroma (Cline and Dabney-Smith 2008; Kessler and Schnell 2009; Jarvis 2008; Schleiff and Becker 2011). This extension was initially thought to be essential and sufficient for targeting and for subsequent translocation of the protein into the organelle (Bionda et al. 2010). Transit peptides do not contain specific amino acid sequences (Bruce 2000; Schleiff and Soll 2000), but they are enriched in positively charged and hydroxylated amino acids (Lee et al. 2006, 2008). It is known that transit peptides can adopt a helical structure in lipid- or protein-rich environments (Bruce 2000, 2001), but the functional relevance of this feature was never proven beyond doubt. It was later shown that proper targeting also requires several classes of cytoplasmic factors (Jackson-Constan et al. 2001; Zhang and Glaser 2002; Schwenkert et al. 2011), which are thought to transport the precursor protein to specific *TRANSLOCONS* in the envelope membranes. The chloroplast translocons consist of two complexes, namely the Translocon of the Outer or Inner envelope membrane of Chloroplasts (TOC or TIC; Oreb et al. 2008; Schleiff and Becker 2011; Sommer and Schleiff 2009). Translocation across the envelope membranes also requires the input of energy which is provided by stromal chaperones (Soll and Schleiff 2004; Li and Chiu 2010).

Several observations suggest that additional import systems exist. Large-scale proteomic approaches identified a number of chloroplast proteins that do not possess a typical transit peptide (e.g., Kleffman et al. 2004) but nevertheless are translocated into chloroplasts. Consequently, these signal sequences were termed *NON-CANONICAL TRANSIT PEPTIDES* (Miras et al. 2002) and are neither cleaved after translocation nor recognized by the classical receptors of the TOC complex (Miras et al. 2007). In addition, evidence for a vesicle-mediated transport of proteins from the Endoplasmic Reticulum (ER) to the chloroplast surface emerged (Villarejo et al. 2005) and translocation of some outer membrane proteins was shown to be dependent on the protein AKR2a (Bae et al. 2008). Evidence accumulated that variations of the general translocation mechanism exist in other plastid types (Bräutigam and Weber 2009) and for specific precursor proteins like

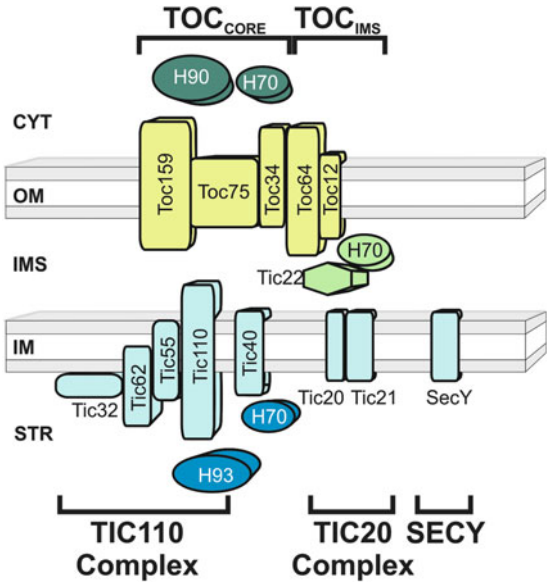


Fig. 1 The chloroplast translocon. The components of the Translocon of the Outer or Inner envelope membrane of Chloroplasts (TOC and TIC) are shown. The components are named according to their localization and the molecular weight of the first identified homolog. The simplified model is not in scale with respect to the size of the proteins or the number of components within one complex. For further details, see text

the NADPH:protochlorophyllide oxidoreductase A (Becker et al. 2005; Reinbothe et al. 2004). Since our understanding of these alternative pathways and the components involved is rather limited, this article will focus on the most common import pathway. We describe the components of the plastid translocon and the bacterial translocation systems before discussing how the more complex plastid translocon may have evolved.

The General Protein Translocon in Land Plants

As mentioned above, the general translocation process across the envelope membranes of plastids involves the two complexes TOC and TIC (Fig. 1). It is not yet certain whether these complexes act in concert or independent of each other. However, attempts to isolate them revealed complexes that are physically separate of each other (Schleiff et al. 2003; Ladig et al. 2011; Scott and Theg 1996; Kikuchi et al. 2009; Kuchler et al. 2002; Caliebe et al. 1997). Unlike in mitochondria, both chloroplast envelope translocons possess their individual energizing subunits: imsHsp70/Toc12 for the outer envelope (Becker et al. 2004; Ruprecht et al. 2010) and a stromal chaperone system consisting of stHsp93/stHsp70/Tic40 for the inner envelope (Kovacheva et al. 2005, 2007; Su and Li 2010; Shi and Theg 2010). In addition, an accumulation of stromal proteins in the InterMembrane Space (IMS) was observed under certain environmental conditions (Hirohashi et al. 2001). All of these factors seem to argue for autonomous action of these complexes.

On the other hand, chemical cross-linking and precursor protein trapping experiments led to the identification of protein complexes including components of both TOC and TIC (Hirohashi et al. 2001; Nielsen et al. 1997; Akita et al. 1997). In addition, it became clear that proteins destined for the inner envelope membrane are translocated first into the stroma from where they are subsequently inserted into the inner membrane (Lubeck et al. 1997; Li and Schnell 2006). This suggests that the IMS-localized translocation system in chloroplasts may be less complex than that found in mitochondria, where inner membrane proteins are inserted from the IMS side (Schleiff and Becker 2011). However, for the evolution of an efficient import system, the two translocons need to be coordinated. Whether that involves direct physical interaction or not remains to be shown.

TOC and TIC (Fig. 1) are composed of three different classes of proteins, namely (1) proteins acting as receptors which directly interact with precursor proteins, (2) proteins forming the required channels for the transfer of the precursor proteins across the membrane, and (3) proteins providing energy for the translocation event or regulating this process in response to various environmental conditions. The receptors of TOC are the two GTPases Toc34 and Toc159 (and homologs thereof), which will not be discussed here (Jackson-Constan et al. 2001; Oreb et al. 2006). Toc75 is the pore-forming unit spanning the outer envelope (Tranel et al. 1995; Schnell et al. 1994; Hinnah et al. 2002). Two factors, namely Toc64 and Toc12, are currently identified as regulatory components (Qbadou et al. 2007; Sohrt and Soll 2000; Becker et al. 2004, 2005). They are thought to be involved in (1) the recognition of cytoplasmic complexes transporting the precursor proteins to the outer envelope, (2) recruiting chaperones of the IMS (imsHsp70) to the membrane, energizing translocation, and (3) linking stromal REDOX regulation to the translocation process (Becker et al. 2004; Oreb et al. 2008). However, regulatory properties are usually hard to prove and thus the detailed mode of action of these two components remains to be elucidated.

In the inner envelope membrane several complexes appear to participate in protein translocation (Fig. 1). The complex that was first identified contains the protein Tic110 which fulfills three functions by interacting with the precursor proteins, providing a channel for translocation and interacting with stromal chaperones (Lubeck et al. 1996; Schnell et al. 1994; Heins et al. 2002). Tic110 forms a complex with four regulatory components, namely Tic62, Tic55, Tic32, and Tic40. The first three proteins are thought to form a complex regulating the translocation efficiency depending on the chloroplast REDOX state (e.g., Benz et al. 2009). Tic40, however, links the translocon to stromal chaperones involved in the translocation process across the inner envelope (Stahl et al. 1999; Chiu and Li 2008; Chou et al. 2003). A complex composed of Tic20 and Tic21 was also identified (Kouranov et al. 1998; Teng et al. 2006) with Tic20 widely accepted as playing a role in translocation. However, the function of Tic21 is more controversial and evidence suggests that this protein may be involved in the transport of iron rather than proteins (Duy et al. 2007). Tic20 forms a translocation channel (Kovacs-Bogdan et al. 2011) and assembles into a complex migrating at a molecular weight of about 1 MDa (Kikuchi et al. 2009), which strongly suggests the existence of

other yet unknown proteins. In fact, a recent analysis of the inner envelope fraction revealed two chloroplast-encoded proteins, namely Ycf1 and Ycf2 migrating at the same molecular weight as this complex in a native PAGE (Ladig et al. 2011).

Furthermore, an inner membrane-localized SecY was identified (Fig. 1; Skalitzky et al. 2011) suggesting the existence of a further translocation system in this membrane. It has been speculated that Tic110 as well as Tic20 may be involved in the translocation of precursor proteins across the inner envelope membrane, and that SecY may be involved in the insertion of plastome-encoded inner envelope proteins such as Ycf1 and Ycf2 (Ladig et al. 2011). Although these proposed functions are mirroring the function of translocons identified in the mitochondrial inner membrane (Schleiff and Becker 2011), experimental evidence regarding their roles is scarce for chloroplasts. Only recently, Kikuchi and coworkers showed, that Tic20 forms a complex with the hithero unknown TIC components Tic214 (Ycf1), Tic100 and Tic56 that was strongly associated with translocating preproteins (Kikuchi et al. 2013).

The Various Translocation Routes Across Prokaryotic Membranes

While our understanding of protein translocation systems in prokaryotes comes mainly from investigations of proteobacterial species, much thought has been given to the origins and evolution of these translocations systems. It has been suggested that the ancestral mode of transport involves the spontaneous insertion of proteins into the membrane without the help of other proteins (Bohnsack and Schleiff 2010; Pohlschroder et al. 2005; Mirus and Schleiff 2012; Mirus et al. 2010). The subsequent evolution of specialized membrane-embedded transport complexes enhanced the specificity and kinetics of the transport and the insertion of preproteins into the membrane. More sophisticated transport systems may also have become necessary as the preproteins that needed to be transported became more complex ranging from single spanning membrane proteins to membrane-embedded multi-domain enzymes. The central component of the ancestral translocon complex might have been a YidC homolog as it catalyzes the insertion of proteins into the membrane (Fig. 2; van der Laan et al. 2005).

Bacteria also evolved more complex translocation systems as they added a cell wall made of peptidoglycan. This required the export of unfolded and folded proteins across the inner membrane to the bacterial surface. The bacterial translocation machineries SEC and TAT are involved in these processes (du Plessis et al. 2011; Mandon et al. 2009; Robinson and Bolhuis 2001; Rusch and Kendall 2007; Driessen and Nouwen 2008). The SEC translocon mediates the translocation of unfolded proteins and requires at least the two components SecY and SecE to assemble into a functional complex (e.g., Tsukazaki et al. 2008). In contrast, TAT (Twin-Arginine-dependent Translocon) is involved in the translocation of folded

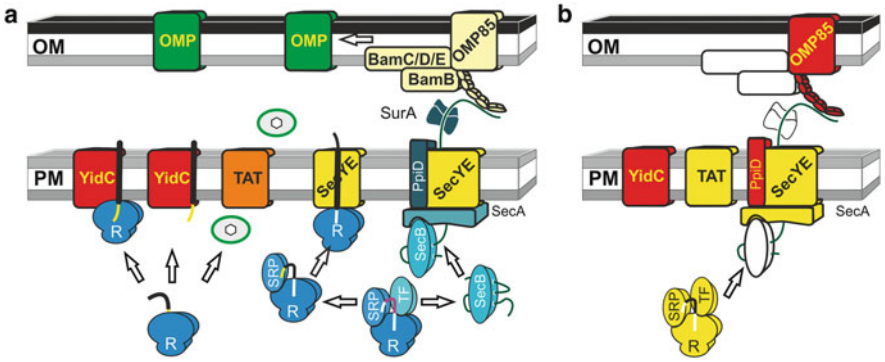


Fig. 2 The translocation systems in prokaryotic membranes. (**a**) The translocation components of the proteobacterial membrane systems are depicted as described in the text. The complexes are divided according to their functional relation and color coding is used to dissect cytoplasmic (*light blue*), plasma membrane (*red*, *orange*, and *yellow*), and outer membrane (*beige* and *green*) complexes. The two components bridging the outer and inner membrane are PpiD and SurA. (**b**) The same dissection as in (**a**) but for cyanobacteria with altered meaning of the *color code*. Here, *red* stands for identified but distinct at sequence level to the one from proteobacteria; *yellow* for existing and rather comparable to proteobacterial sequences; *white* stands for not existing

proteins across the inner membrane and consists of class one type components, namely the receptors TatB and TatC, and a class two type component, the pore-forming unit TatA (Palmer et al. 2005; Natale et al. 2008; Robinson and Bolhuis 2001).

In eukaryotes, the components involved in protein targeting serve three functions: (1) the recognition of the topogenic signal defining the protein as non-cytoplasmic, (2) its targeting to the organelle, and at the same time (3) the targeting to distinct membrane destinations within that organelle. In bacteria, protein targeting has to fulfill only two of these functions. A central cytoplasmic component for (1) is the Srp54 (Signal recognition particle component of 54 kDa) homolog Ffh, which exists ubiquitously in all sequenced bacterial genomes (Grudnik et al. 2009). Ffh is thought to specifically recognize the bacterial secretion signals and thus allows for proteins to be excreted in a SEC-dependent manner.

Gram-negative bacteria are surrounded by a second membrane (Fig. 2) and, consequently, need additional translocation machineries to supply this outer membrane with proteins. It has been shown that additional cytoplasmic factors are involved in the transport of proteins to the membrane surface, e.g., the chaperone-like protein SecB (Driessen et al. 2001) and the Trigger Factor (TF). Trigger factor is a ribosome-associated protein which is involved in the discrimination between SRP- and SecB-dependent pathways (Beck et al. 2000). SecB, a soluble chaperone, might be involved in initial protein folding rather than the transport of proteins to the membrane surface (Ullers et al. 2004). In addition, periplasmic and outer membrane-localized translocation systems are required, which may serve similar functions as those in the cytoplasm and the inner membrane. The periplasmic components involved in targeting towards the outer

membrane are the chaperone-like proteins SurA and Skp (Knowles et al. 2009; Sklar et al. 2007). The notion that the periplasmic system is similar in function to the one in the cytoplasm is supported by the observation that the putative substrate binding grooves of TF and the periplasmic chaperone SurA are structurally similar (Stirling et al. 2006). Furthermore, Skp is thought to serve as a general chaperone for proteins targeted to the periplasm, which is similar to the function of SecB (Sklar et al. 2007). The insertion of outer membrane proteins is catalyzed by Omp85 (Knowles et al. 2009; Schleiff and Becker 2011; Hagan et al. 2011; Löffelhardt et al. 2007), but it has been suggested that outer membrane proteins originally inserted spontaneously into the membrane (Mirus and Schleiff 2012). In *E. coli*, a larger complex has been identified which contains Omp85 (here annotated as BamA) and the proteins BamB, BamC, BamD, and BamE forming the so-called BAM (β-Barrel Assembly Machinery; Hagan et al. 2011).

Chloroplasts originated most likely from the same cyanobacterial ancestor as the *Nostocales* (Deusch et al. 2008; Martin et al. 2002), triggering much interest in translocation in cyanobacteria. Recent work has identified differences between proteo- and cyanobacteria with respect to the existing components involved in protein translocation. In the plasma membrane-localized translocation systems significant differences in the amino acid sequence of the YidC homolog TatA were observed (Bohnsack and Schleiff 2010; Mirus and Schleiff 2012). The proteobacterial YidC proteins contain an additional N-terminal domain compared to the cyanobacterial proteins. For the TAT translocase, the detection of TatB in cyanobacteria is compromised by its similarity to TatA (Bohnsack and Schleiff 2010; Mirus and Schleiff 2012). However, two TatA-like proteins have been detected in cyanobacterial genomes and thereby it is most likely that the cyanobacterial TAT system is comparable to that described in proteobacteria. Two proteins that are involved in the insertion of outer membrane proteins in proteobacteria could not be identified in cyanobacteria, namely SecB and SurA, while for PpiD which is a periplasmic peptidyl prolyl *cis–trans* isomerase involved in the release of proteins from the SEC translocon (Stymest and Klappa 2008) a homologous sequence was identified in cyanobacterial genomes (Bohnsack and Schleiff 2010). The same holds true for the BAM components except for the Omp85 homolog BamA, which is present in both groups. However, the two functional domains of the proteo- and cyanobacterial Omp85 proteins show clear differences. First, their periplasmic POTRA domains show alterations in structure and amino acid sequence (Koenig et al. 2010) maybe because different periplasmic chaperones are present in these bacteria. Second, the β-barrel-shaped pore domains of the proteins have different properties in proteo- and cyanobacteria (Bredemeier et al. 2007). Thus, although the Omp85 proteins share a common ancestor (Bredemeier et al. 2007; Moslavac et al. 2005), they are not entirely functionally orthologous (e.g., Wunder et al. 2007) and have undergone diversification. For example, in the filamentous cyanobacterium *Anabaena* sp. PCC 7120 more than one Omp85 homolog was found (Nicolaisen et al. 2009).

The Evolution of the Targeting and Translocation Mechanisms into Plastids

The Possible Origin of the Targeting Mechanism

It is obviously important to understand by which mechanism the first cytoplasmically produced proteins were imported into the newly aquired endosymbionts. To answer this question it is crucial to know at which evolutionary state of the host cell the uptake of the cyanobacteria occurred. Most likely the last common eukaryotic ancestor contained already an enveloped nucleus, an endomembrane system (Bohnsack and Schleiff 2010; Field and Dacks 2009), and a mitochondrion (van der Giezen and Tovar 2005). Thus, the ancestral eukaryotic cells that took up the cyanobacteria were most likely already equipped with simple mechanisms for the targeting and transport of proteins to cellular compartments. These cells may have either utilized vesicular transport or exploited targeting systems involving soluble chaperone-like factors, as have been shown to exist for mitochondria. Therefore, the mechanism of transport into the precursors of chloroplasts had to be integrated into the existing system and might have been partially derived from that. While the gene transfer from bacteria to the host nucleus is a rather common event, the evolution of a protein targeting and translocation mechanism was probably the bottleneck for the transition from an endosymbiotic bacterium to an integrated organelle (Stegemann and Bock 2006; Stegemann et al. 2003).

A possible scenario involves the "hijacking" of the vesicular system for the transport of the first cytoplasmically synthesized "plastid" proteins. Especially thylakoid proteins, which are membrane-anchored and possess numerous hydrophobic transmembrane segments that could serve as ER targeting signals, may have utilized the vesicle transport pathway. In addition, proteins with secretion signals that do not differ significantly between prokaryotes and eukaryotes might have been transported via in this manner. In fact, some subunits of photosystem I in *Paulinella chromatophora* seem to be targeted to the chloroplast via the endomembrane system (Nowack and Grossman 2012), which provides support for such a scenario. Vesicles would have to bud from the ER, travel to the chloroplast, and fuse with its outer membrane. However, this does not solve the problem of how proteins are subsequently released from the outer membrane and inserted into the inner membrane. Overall, an interorganellar vesicular transport system appears to be unlikely to have existed since the protoplastid probably still was surrounded by a peptidoglycan layer, which would render the transport of a larger vesicle rather difficult. Our ideas about early plastids are based on the fact that the primordial plastid in glaucophytes (Steiner and Löffelhardt 2005) still contains this peptidoglycan layer. Thus, the chloroplast would need to be able to provide enough energy to extract the delivered proteins from the outer membrane into the intermembrane space, in a process similar to that for mis-folded ER membrane proteins (Wolf and Stolz 2012). If such a system indeed evolved, the proteins might have utilized the SEC translocon in a retrograde manner to insert proteins into the inner membrane.

At this state of evolution, the protoplastid may still have had continuous thylakoid and inner envelope membranes (Zak et al. 2001), which means the proteins could have entered the thylakoid membranes by lateral diffusion from the inner envelope membrane. Later in the evolution of chloroplasts this lateral diffusion was probably replaced by vesicular transport between the two membranes as connections between them were severed (Vothknecht and Soll 2005). This idea is supported by the identification of proteins involved in vesicle transport in proteomic studies of chloroplast envelope fractions (Bräutigam and Weber 2009). Another example showing that vesicle transport between the endomembrane system and symbiont is possible comes from legumes in which the transition of the nodule bacteria to bacteroids depends on a specific vesicle transport via the secretory pathway to the peribacteroid membrane (van de Velde et al. 2010; Verma and Hong 1996; Mergaert et al. 2003).

However, several observations challenge the scenario of an exclusively vesicle-based translocation system as the origin for the translocation mechanism in chloroplasts. The previously described scenario requires that a vesicle fusion system in the outer membrane had evolved prior to the first retrograde translocation. In addition, specificity of the vesicle transport system is necessary to avoid targeting of the thylakoid proteins to the wrong location. Assuming those requirements were met, the thylakoid proteins would be localized in the outer membrane after vesicle fusion. If this scenario is correct one would expect to find evidence for multiple signals, namely for the co-translational import into the ER lumen (Hiss and Schneider 2009), for the subsequent sorting to the chloroplast (Jürgens 2004), *and* for the translocation across the inner membrane of the organelle. This system would have worked well for proteins with segments with a high hydrophobicity (see above). However, many of the highly hydrophobic proteins are plastome encoded (e.g., Allen 2003).

A second scenario proposes that stromal proteins, which were synthesized on cytoplasmic ribosomes, possessed some kind of targeting signal and a translocation apparatus allowed their transport across the outer and inner envelope membrane. Several lines of evidence support the idea that this was indeed the original system for protein translocation, which is very similar to what exists now. In the following section we will describe the origin of this system.

The Targeting Signals and the Targeting Complexes

The evolution of chloroplasts after the transfer of the vast majority of the cyanobacterial genome to the host nucleus required the development of targeting and translocation systems for the retrograde transport (Cavalier-Smith 2003). Most of the nuclear-encoded stromal proteins known today are synthesized as precursor proteins possessing an N-terminal targeting sequence, the so-called transit peptide (Schleiff and Becker 2011). As discussed earlier, several different transport routes are known to exist today (Section “The Complexity of Protein Translocation into

Chloroplasts"), but most likely the transit peptide-mediated pathway represents the original mode for translocation since functional transit peptides already existed in glaucophyte muroplasts (cyanelles; Steiner and Löffelhardt 2005).

The evolution of the N-terminal signals had to go hand in hand with the evolution of the translocation machineries, as they are the decoders of these signals. It is believed that the chloroplast translocation channel Toc75 originated from Omp85 (Löffelhardt et al. 2007; Schleiff and Becker 2011). The latter is a β-barrel-shaped outer membrane protein assembly factor (see section "The Various Translocation Routes Across Prokaryotic Membranes"), which recognizes proteins containing a C-terminal phenylalanine or tryptophan residue as substrate (see Box 1; Scheufler et al. 2000). Remarkably, targeting signals of muroplast proteins of the glaucophyte *Cyanophora paradoxa* contain an essential phenylalanine close to their N-terminus (Steiner and Löffelhardt 2005) which drives the interaction of the precursor proteins with proteins of the Omp85 family (Wunder et al. 2007).

Box 1: The Evolution of Toc75 as Primitive TOC Translocon from the Ancestral Omp85

Nowadays two distinct plastidic Toc75 families exist in land plants, namely one for the insertion of outer membrane proteins and one for protein translocation into the organelle. The first has the original function of Omp85 proteins. Thus, a central question of the evolutionary development of the protein translocon in the outer envelope membrane concerns the timing of the duplication and the functional diversification of this ancient Omp85 and the time point at which the change of its topology occurred. In cyanobacteria (Figure top, light yellow background) the POTRA domain was exposed to the periplasm and the Omp85 served as β-barrel protein "insertase" (in dark green) for the outer membrane acting from the periplasmic surface (lateral arrow), recognizing substrates by their C-terminal phenylalanine motive (F; Struyve et al. 1991; Hagan et al. 2011; Knowles et al. 2009; Ruiz et al. 2006). The birth of the ancient translocon rather coincided with the transfer of its encoding gene to the host nucleus than when still being encoded on the symbiontic genome (yellow). Further, the initial driving force for the development of an ancient "TOC" was probably the need for the import of host proteins into the symbiont's inner membrane for the extraction of energy-rich metabolites than to establish a reimport of the first symbiontic proteins (vertical arrows) which had already been integrated into the host genome. This view is tempting, especially since porines for metabolite transport across the outer membrane already existed in the symbiont. At stage, there is no information on whether the ancestral nuclear-encoded Omp85 (light green) was imported with a topology exposing the POTRA to the periplasm (state I, right) or whether the inversion of the Omp85 exposing the POTRA to the

(continued)

Box 1. (continued)

cytoplasm coincided with the retrograde transport of this protein (state I, left). This inversion might have been simply a consequence of the targeting to the opposite side of the membrane (the cytoplasmic surface). However, the inversion has to be considered as an early event as it exposed the substrate recognizing protein surface into the compartment where the precursor proteins were delivered to. This view is consistent with the observed topology of Omp85 in diatoms (Bullmann et al. 2010) and plants (Sommer et al. 2011).

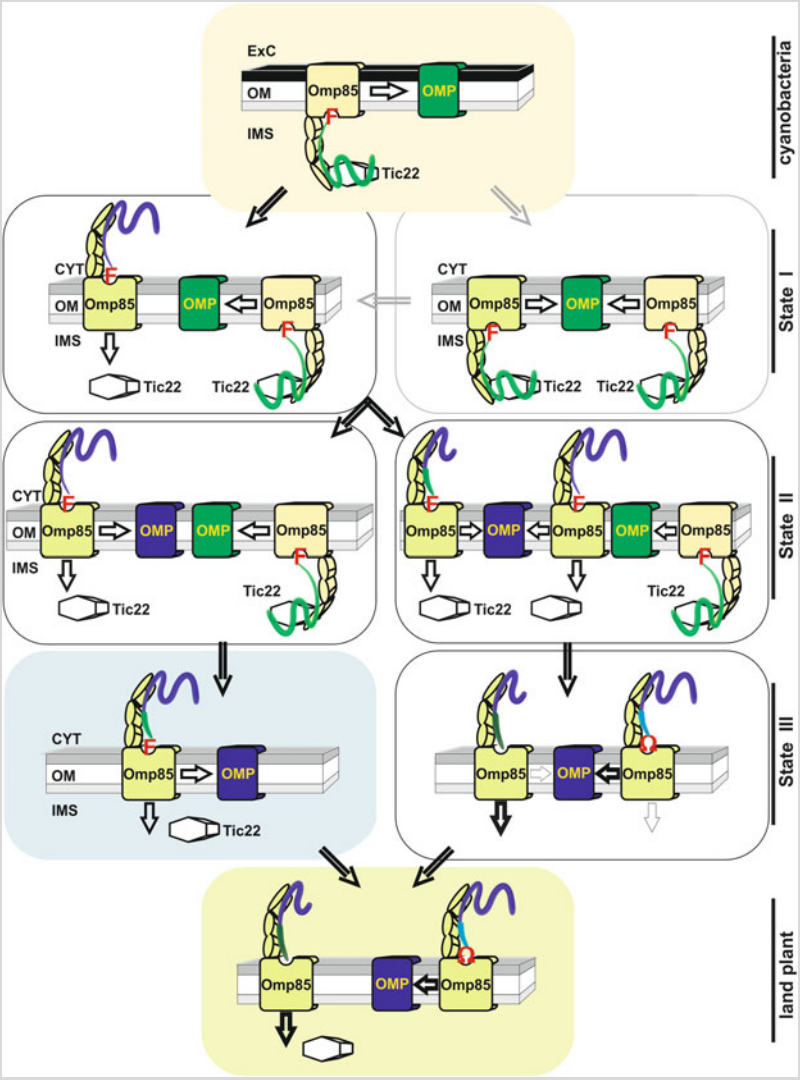
Two different possible scenarios for the subsequent evolution of the primitive TOC during the transition can be imagined. On the one hand (right side), the nuclear-encoded Omp85 duplicated first, with both proteins functioning in protein translocation and protein insertion of the now nuclear-encoded outer membrane proteins (blue) into the membrane (state II, right). In the course of evolution all outer membrane proteins have been transferred into the nuclear genome resulting in the loss of the bacterial-encoded Omp85 protein (state III). At this stage a certain functional diversification might have taken place as the two different Toc75 coevolved independently from each other along with specific topogenic signals (transit peptides for preprotein import and aromatic amino acids (Ω) for OMPs) for both pathways (state III, right). Alternatively, the nuclear-encoded Omp85 might have served both the translocase for "inner" organelle proteins and the insertase function for the relocated OMPs (state II, left). Again, the nuclear transfer of all outer membrane coding genes made the symbiont-encoded Omp85 gradually dispensable (state III), while the one nuclear-encoded protein fulfilled the two functions. Thus, gene duplication and diversification into importer (Toc75-III) and independent insertase (Toc75-V) as known for green algae and plants occurred rather late by this scenario, and at this stage the diversification of the signals continued towards transit peptides not depending on the N-terminal phenylalanine (bottom).

The latter mode of evolution, however, appears to be more likely. On the one hand only a single Omp85 was found in the glaucophyte *Cyanophora paradoxa*, which harbors the most primitive plastids we know today, as well as in cryptophytes and diatoms (state III, left; Bullmann et al. 2010; Price et al. 2012). These Omp85 proteins still rely on the primordial phenylalanine motive within the transit peptides of nuclear-encoded plastid proteins as targeting signal (Wunder et al. 2007; Kilian and Kroth 2005; Gould et al. 2006; Steiner et al. 2005). Remarkably, in land plants Toc75-III and Toc75-V do not share the same import pathway (Inoue and Potter 2004). Toc75-III is the only outer envelope protein of chloroplasts from land plants possessing a cleavable transit peptide, followed by a glycine-rich amino acid stretch (Tranel et al. 1995). Both the Toc75-V-like protein of

(continued)

Box 1. (continued)

Chlamydomonas reinhardtii (Fig. 3) and the general Omp85 in diatoms (Bullmann et al. 2010) contain the typical poly-glycine stretch. This strongly suggests that the transit peptide and the poly-glycine motif occurred rather early in evolution and have been subsequently lost in the Toc75-V branch.



Similarly, most transit peptides of red algae (e.g., of *Cyanidioschyzon merolae*; Patron and Waller 2007) and transit peptides of proteins for transport into the stroma of secondary plastids which have derived from red algae (Armbrust et al. 2004; van Dooren et al. 2001; Kilian and Kroth 2005; Harb et al. 2004; Ralph et al. 2004; Grosche et al. 2013) contain a phenylalanine in their very N-terminus as well. Hence, this property of the ancestral Omp85 to recognize the phenylalanine may have dictated the initial evolution of transit peptides (Mirus and Schleiff 2012).

Proteins without a targeting signal present in the cytoplasm of the host cell had to be furnished with a signal. One mode discussed for the acquisition of these signals includes exon shuffling (Bruce 2001). Support for this idea comes from a recent analysis of nuclear-encoded mitochondrial ribosomal subunits in rice and *Arabidopsis thaliana*. It was shown that 19 of 30 proteins have an N-terminal extension which is not present in the bacterial ancestor, and interestingly the extension is encoded by an individual exon (Bonen and Calixte 2006). Moreover, when analyzing the genomic regions coding for the first 100 amino acids of all *A. thaliana* proteins with defined but different intracellular localizations (according, e.g., to the Plant Proteome Database), we did not find a specific enrichment of intervening sequences in genes encoding plastid preproteins. But when those genes possessed such an intron it was often located in the cleavage site between the transit peptide and the mature domain on protein level (Mirus and Schleiff 2012), which supports the exon-shuffling hypothesis.

A second mechanism for transit peptide evolution could be the recycling of the 5′-UTR of the respective bacterial genes, as shown by the analysis of the O-acetylserine (thiol)-lyases (Rolland et al. 1993). Although this idea is very appealing, the probability that the 5′-UTR of a bacterial gene can be adapted to serve as a transit peptide with the specific properties required for retrograde transport is low. Nevertheless, currently there is not enough evidence to exclude either model and maybe several different processes were involved.

Transit peptides have to remain largely unfolded for proper targeting. Thus, a system evolved in which they associated with molecular chaperones, especially Hsp70 (Zhang and Glaser 2002) and Hsp90-type proteins (Qbadou et al. 2006) which are known to participate in protein folding. Both chaperones are of prokaryotic origin and thus must have been present in the host cell (Mirus and Schleiff 2009; Gupta 1995; Gupta and Golding 1993). This initial assembly may have led to the evolution of targeting complexes, which in many cases contain these two types of molecular chaperones (Schleiff and Becker 2011). In the course of evolution, the importance of chaperones may have shifted from targeting to the maintenance of an unfolded import-competent state of the precursor proteins (Ruprecht et al. 2010), at least for Hsp70-guided proteins. In addition, control mechanisms for precursor protein degradation involving a specific Hsp70 (Hsc70-IV in *Arabidopsis thaliana*) evolved (Lee et al. 2009). Thus, the main function of Hsp70 in the transport remained in the regulation of the folding state of the preproteins. However, some

limited evidence exists that Hsp90 proteins might be actively involved in protein targeting as well (Qbadou et al. 2006; Fellerer et al. 2011).

The Evolution of the Outer Envelope Translocon

The TOC is composed of Toc64, Toc34, Toc159, Toc12, and Toc75 (Fig. 1; Oreb et al. 2008; Schleiff and Becker 2011). As mentioned above (Section "The General Protein Translocon in Land Plants") the central, pore-forming unit Toc75 belongs to the Omp85 protein family (Fig. 3; Löffelhardt et al. 2007) and is closely related to cyanobacterial Omp85 proteins (Bredemeier et al. 2007; Yusa et al. 2008). Omp85 proteins generally are composed of a number of N-terminal, so-called POlypeptide–TRansport-Associated (POTRA) domains (Schleiff and Becker 2011; Koenig et al. 2010; Schleiff et al. 2011) and a C-terminal 16-stranded β-barrel domain (e.g., Jacob-Dubuisson et al. 2009). The C-terminal domain of the cyanobacterial Omp85 has a pore dimension suitable for the translocation of unfolded polypeptides as exemplified, e.g., for the protein of *Anabaena* sp. (Bredemeier et al. 2007). The POTRA domains are possibly involved in the formation of homo- and hetero-oligomeric complexes (Bredemeier et al. 2007; Ertel et al. 2005). In bacteria, this domain is exposed to the periplasm and interacts with periplasmic chaperones (Section "The Various Translocation Routes Across Prokaryotic Membranes"; Ieva et al. 2011; Tripp et al. 2012).

In plastids, the Omp85 homolog Toc75-III is the only outer envelope protein with a transit peptide (Tranel and Keegstra 1996; Inoue et al. 2001). The transit peptide is bipartite, where the first part is cleaved in the stroma and the second part in the inter membrane space by a type I signal peptidase (Inoue et al. 2005). The second portion contains a poly-glycine stretch which is important for proper translocation (Inoue and Keegstra 2003; Baldwin and Inoue 2006). This signal, however, is specific to the Toc75 protein forming the major translocation channel in land plants and does not exist in Toc75-like proteins involved in the insertion of other outer membrane proteins (Fig. 3, section "The Evolution of the Outer Envelope Protein Insertion Machinery"). In addition, a C-terminal motif (FGERF) of unknown function is characteristic for all Toc75 proteins involved in translocation (Fig. 3), while an additional glycine-rich segment exists in the C-terminal portion of all Toc75 proteins. Remarkably, this poly-glycine stretch also exists in the C-terminus of the cyanobacterial Omp85 proteins as well as the C-terminal GERF motif (Fig. 3). Thus, it is likely that these two motifs were inherited from the cyanobacterial ancestors and might be involved in the insertion of the Toc75 proteins into the membrane as shown for bacterial outer membrane proteins (Struyve et al. 1991).

Remarkably, all sequences of Toc75 proteins most likely involved in insertion of proteins into the membrane (Section "The Evolution of the Outer Envelope Protein Insertion Machinery") do not contain a C-terminal phenylalanine with the exception of the proteins found in some Chlorophyta (Fig. 3). In addition, only the Toc75

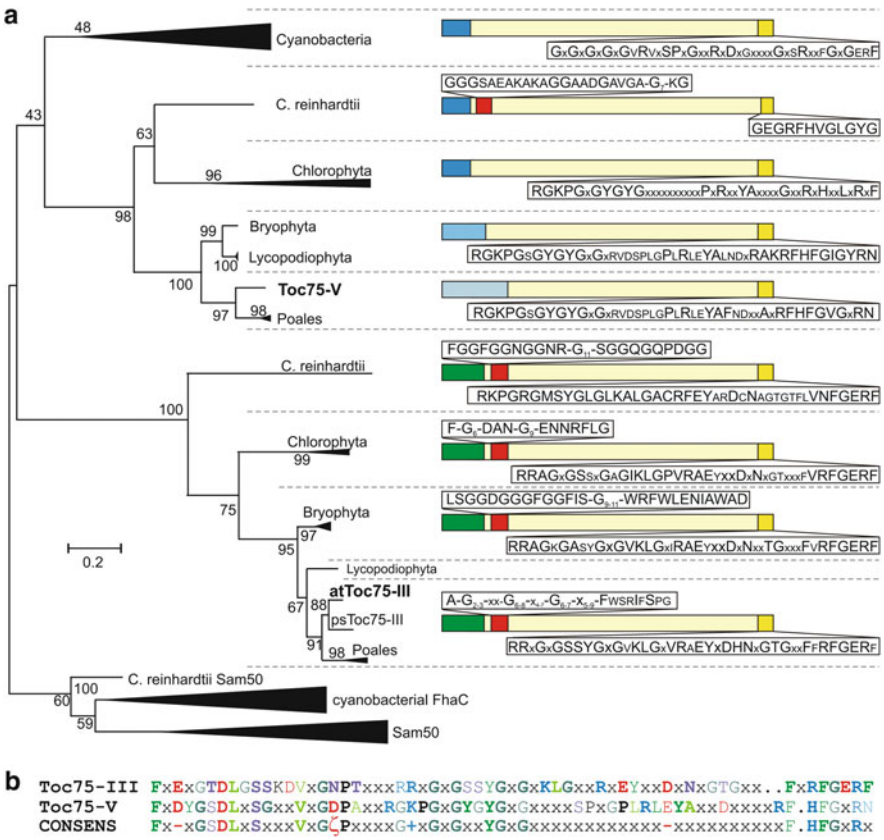


Fig. 3 The phylogenetic diversification of the Toc75 family. (**a**) The phylogeny of Toc75 protein sequences listed in Table 1 is shown on the left, where clades are compressed for clarity. The sequences of the Omp85-TpsB transporter superfamily protein FhaC and of the mitochondrial Omp85 homolog Sam50 are used as outgroup. For each clade the properties of the sequences were analyzed with respect to the existence of a signal sequence analyzed by SignalP (Petersen et al. 2011) for cyanobacterial and Toc75-V proteins and with TargetP (Emanuelsson et al. 2000). Signals identified with SignalP are indicated by a *blue bar* and signals identified by TargetP are indicated by a *green bar*. The sequences were further analyzed with respect to the poly-glycine region behind the signal sequence. Its existence is indicated by a *red* region and the consensus sequence is given on *top*. The C-terminus of the proteins of each clade was analyzed with respect to the existence of a glycine-enriched portion and the amino acid composition of the extreme C-terminus. The consensus sequence is given *underneath the bar*. (**b**) The consensus of the extreme C-terminus of the analyzed Toc75-V proteins (involved in protein insertion into the outer membrane), the Toc75-III proteins (involved in translocation across the envelope membrane), as well as the overall consensus is given

for membrane protein insertion in *Chlamydomonas reinhardtii* contains an N-terminal poly-glycine stretch (Fig. 3). This strongly suggests that the targeting signal evolved very early and was subsequently lost in the proteins of this clade (Fig. 3).

In addition to the evolution of a transit peptide replacing the secretion signal of the bacterial Omp85, the topology of the Toc75 was changed such that the POTRA domains are now exposed to the cytoplasm (see Box 1; Bullmann et al. 2010; Sommer et al. 2011). This change in orientation was crucial for the evolution of the preprotein targeting signals (Section "The Possible Origin of the Targeting Mechanism"), because only an inverted topology allowed the new evolved Toc75 to interact with cytoplasmic precursor proteins possibly via the phenylalanine motif. It is likely that Toc75 alone was originally the functional basic translocon, as both the chloroplast and the cyanobacterial POTRA domains are able to interact with transit peptide containing precursor proteins (Ertel et al. 2005). Even more though, the interaction between eukaryotic precursor proteins and the cyanobacterial Omp85 was found to be enhanced in the presence of a phenylalanine at the N-terminus of the transit peptide (Wunder et al. 2007). This suggests that the early translocon served both in protein translocation and outer membrane protein insertion. This dual function might still be found in Toc75 in glaucophytes, rhodophytes, and complex plastids derived from red algae. In the course of evolution the specificity might have changed as multiple Toc75 homologs emerged (see Box 1), possibly leading to the diversification of the transit peptides away from the phenylalanine requirement in the green algae and plants, and accompanied by the evolution of highly specific receptors (such as Toc34 and Toc159 for the TOC complex). The need for the latter might be explained by a higher demand for efficient import or regulatory circuits to suffice the complexity of the photosystems and antenna complexes.

The need for enhanced efficiency and selectivity of precursor protein recognition and translocation led presumably to the evolution of additional receptor components with a clear eukaryotic origin as indicated by their helical transmembrane domains. Toc159 and Toc34 are dimerizing GTPases and belong to the class of TRAnslation FACtor-related (TRAFAC) G proteins (Leipe et al. 2002). Thus, the two G(TPase) domains share a high degree of similarity (Oreb et al. 2008), but the remaining protein modules are distinct. Toc34 has a single C-terminal transmembrane helix, while Toc159 has a C-terminal 52 kDa M(embrane) domain and an additional N-terminal A(cidic) domain, the latter possibly of disordered nature (Richardson et al. 2009). However, based on the similarity of the G-domain one can assume that (1) Toc159 and Toc34 originated from a common ancestral GTPase, and that (2) gene duplication might have led to two distinct receptors (Oreb et al. 2008). The sequencing of the genome of *Cyanophora paradoxa* led to the discovery of two Toc34, but no Toc159 homolog. As the annotation of Toc34 is only based on the homology of the G-domain, it is not clear whether these proteins are real Toc34 homologs (Price et al. 2012). It was previously suggested that both receptors exist in red and green algae (Kalanon and McFadden 2008); however, especially the putative homolog of Toc159 is clearly distinct from the proteins in land plants (Fig. 4). Interestingly, it was noticed that the Toc34 sequences from red and green algae, especially from *Chlamydomonas reinhardtii*, contain an acidic extension at the N-terminus, which might be a rudimentary A-domain (Kalanon and McFadden 2008). This evidence along with the identification of putative Toc34

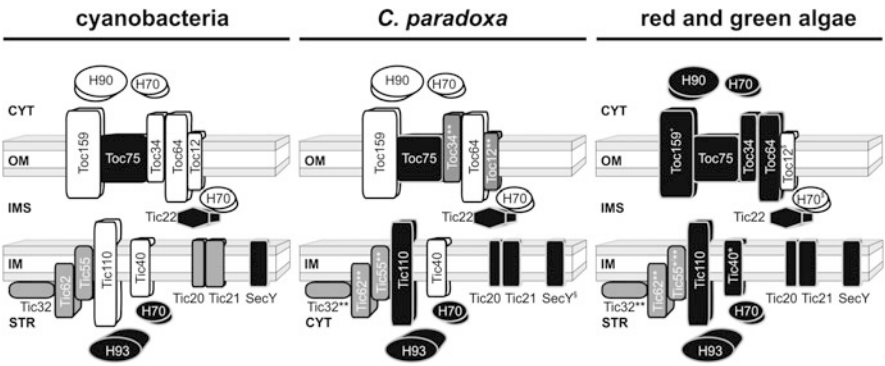


Fig. 4 The evolution of the general translocation path into chloroplasts. The evolutionary development of the translocation components in the chloroplast envelope membranes is depicted based on the model presented in Fig. 1. *On the left* the proteins already existing in cyanobacteria are shown. *Black* indicates proteins with similar function and *gray* proteins with sequence relation but a (putative) function distinct from protein transport. *White* stands for factors for which no homologs are found in cyanobacteria. In the middle the putative translocon composition in *C. paradoxa* (Price et al. 2012) is presented. *Black* and *white color* has a similar meaning as on the *left side*. *Gray* indicates factors where a sequence was identified but with low certainty. *On the right*, the composition in red and green algae is indicated with the same color code as shown in the *middle*. The assignment is based on the data provided by Kalanon and McFadden (2008). *Asterisk* marks proteins identified in green algae only, *double asterisks* for uncertain assignment of the identified factor as transport component, *superscript plus* for uncertain assignment in red algae, *superscript section sign* for plastome encoded, and *superscript dollar sign* for not analyzed

genes in *Cyanophora paradoxa* makes it tempting to speculate that this protein indeed was the ancestral TOC receptor GTPase (Fig. 4).

Toc64 has soluble domains in both the cytoplasm and the intermembrane space separated by transmembrane helices (Qbadou et al. 2007). The cytoplasmic domain is a clamp-type TPR domain involved in Hsp90 recognition (e.g., Qbadou et al. 2006; Mirus and Schleiff 2009; Mirus et al. 2009), while the IMS domain represents a silenced amidase domain (Qbadou et al. 2007; Sohrt and Soll 2000). The protein is a clear eukaryotic invention, does not exist in cyanobacteria or *Cyanophora paradoxa* (Fig. 4), and originated most likely from the TPR domain of HIP proteins (Schlegel et al. 2007). Remarkably, in land plants a paralog of the chloroplast-localized Toc64 is present in mitochondria, where it fulfills a comparable function (Chew et al. 2004). Analysis based on homology models of the TPR domain of the Toc64 paralogs uncovered that amino acids discriminating the differently located proteins are almost exclusively on the surface and not involved in chaperone recognition (Mirus et al. 2009). Thus, the Hsp90 recognizing properties are conserved between the two proteins, while the region recognizing other TOM or TOC components has coevolved with the different interaction partners in the respective complexes.

Besides Toc64 two additional components of the TOC translocon face the IMS, namely imsHsp70 and Toc12 (Marshall et al. 1990; Becker et al. 2004). The Hsp70

Table 1 Genes used for phylogenetic analysis presented in Fig. 3

Gene number	Species	Gene number	Species
gil18419973	*Arabidopsis thaliana*	gil168060103	*Physcomitrella patens*
gil326508096	*Hordeum vulgare*	gil167998366	*Physcomitrella patens*
gil212275868	*Zea mays*	gil168007905	*Physcomitrella patens*
gil302769300	*Selaginella moellendorffii*	gil302769253	*Selaginella moellendorffii*
gil302764336	*Selaginella moellendorffii*	gil7267551	*Arabidopsis thaliana*
gil167997731	*Physcomitrella patens*	gil303281122	*Micromonas pusilla* CCMP1545
Cre02. g122700	*Chlamydomonas reinhardtii*	gil145343539	*Ostreococcus lucimarinus* CCE9901
gil303284771	*Micromonas pusilla* CCMP1545	gil168016338	*Physcomitrella patens*
gil255086251	*Micromonas* sp. RCC299	gil17229151	*Anabaena* sp. PCC 7120
gil242091043	*Sorghum bicolor*	gil166366836	*Microcystis aeruginosa* NIES-843
gil42565501	*Arabidopsis thaliana*	gil17232608	*Anabaena* sp. PCC 7120
gil7635461	*Arabidopsis thaliana*	gil300864043	*Oscillatoria* sp. PCC 6506
gil300867385	*Oscillatoria* sp. PCC 6506	gil17228681	*Anabaena* sp. PCC 7120
gil166366662	*Microcystis aeruginosa*	gil300869127	*Oscillatoria* sp. PCC 6506
gil159027880	*Microcystis aeruginosa*	gil126656575	*Cyanothece* sp. CCY0110
gil126660921	*Cyanothece* sp. CCY0110	gil126658571	*Cyanothece* sp. CCY0110
gil16330104	*Synechocystis* sp. PCC 6803	gil126660764	*Cyanothece* sp. CCY0110
gil17229761	*Anabaena* sp. PCC 7120	gil16332242	*Synechocystis* sp. PCC6803
gil294979844	*Anabaena* sp. PCC 7120	gil126661315	*Cyanothece* sp. CCY0110
gil170076946	*Synechococcus* sp. PCC 7002	gil166365226	*Microcystis aeruginosa* NIES-843
gil37521432	*Gloeobacter violaceus* PCC 7421	gil126661159	*Cyanothece* sp. CCY0110
gil17232385	*Anabaena* sp. PCC 7120	CMJ202C	*Cyanidioschyzon merolae*
gil17227571	*Anabaena* sp. PCC 7120	gil303279551	*Micromonas pusilla* CCMP1545
gil300863869	*Oscillatoria* sp. PCC 6506	gil145347803	*Ostreococcus lucimarinus* CCE9901
gil78779820	*Prochlorococcus marinus*	Cre06. g308900	*Chlamydomonas reinhardtii*
gil145352166	*Ostreococcus lucimarinus* CCE9901	gil18414910	*Arabidopsis thaliana*
Cre03. g175200	*Chlamydomonas reinhardtii*	gil15228433	*Arabidopsis thaliana*
gil15232625	*Arabidopsis thaliana*	CMO061C	*Cyanidioschyzon merolae*
gil242041439	*Sorghum bicolor*	gil242054867	*Sorghum bicolor*
gil75221490	*Pisum sativum*	gil302768034	*Selaginella moellendorffii*
gil242046068	*Sorghum bicolor*	gil168005826	*Physcomitrella patens*

is of eukaryotic (Schnell et al. 1994) and clearly not of cyanobacterial origin. Toc12 is related to bacterial DnaJ proteins. It has neither been identified in the green alga *Chlamydomonas reinhardtii* nor in the moss *Physcomitrella patens* (Kalanon and McFadden 2008), but was annotated in *Cyanophora paradoxa*. However, DnaJ-type proteins are usually not localized in the outer membrane of cyanobacteria.

Since their function has to be investigated in relation to Hsp70, which is of eukaryotic origin, both proteins presumably coevolved.

Given the origin of the different proteins and their identification in *Cyanophora paradoxa*, it is believed that Omp85 was at the center of the evolution of the outer envelope translocon. One essential step was the inversion of the topology (Bullmann et al. 2010; Sommer et al. 2011; Steiner and Löffelhardt 2005). The POTRA domains are known to associate with transit peptides (Ertel et al. 2005) and chaperones (Sklar et al. 2007; Tripp et al. 2012) and hence probably serve as the initial receptor for protein translocation. During evolution the receptor proteins Toc64, Toc34, and Toc159 were added. Based on the genome of *Cyanophora paradoxa* and the existence of the charged N-terminus it appears likely that Toc34 evolved prior to Toc159. The occurrence of the additional receptor units restricted the relevance of the Toc75 protein to its pore-forming properties, and thus, Toc75-like proteins without a POTRA domain could evolve (Schleiff and Becker 2011). It is likely that the other two components Toc12 and imsHsp70 evolved in concert with Toc64, which is involved in the intermembrane space complex, but this remains to be proven.

The Evolution of the Inner Envelope Translocon

The TIC translocon is largely of cyanobacterial origin (Fig. 4), but evolved probably through massive gene recycling rather than through the use of preexisting protein translocon components, such as SecYE, TAT, or YidC type (Fig. 2). Solely for the recently identified inner envelope-localized SecY (Skalitzky et al. 2011) and for the intermembrane space-localized Tic22 has a clear prokaryotic origin has been demonstrated (Fulda et al. 2002; Tripp et al. 2012). Interestingly, the cyanobacterial Tic22 is localized in the thylakoid lumen as well as in the periplasm, which might point towards the existence of a periplasma–lumen connection as previously discussed (e.g., Spence et al. 2003). Tic22 interacts with Toc64 and Toc12 (Qbadou et al. 2007; Becker et al. 2004) and with the pore-forming inner envelope protein Tic20 (Kouranov et al. 1998). Tic20 belongs to the PReprotein and Amino acid Transporter (PRAT) family and shows sequence similarity to the cyanobacterial LivH proteins (Mirus and Schleiff 2012; Bodyl et al. 2010). Tic20 interacts with Tic21, which is of cyanobacterial origin as well (Lv et al. 2009), but its role in protein translocation remains under debate.

The second translocon component suggested to be a pore-forming protein, Tic110, is not of cyanobacterial origin (Fig. 3). Nevertheless, Tic110 seems to exist in glaucophytes as determined by Western blotting in *Cyanophora paradoxa* (Yusa et al. 2008) and was also found in the recently sequenced genome (Price et al. 2012). Thus, Tic110 has to be considered a very early eukaryotic invention. This strongly might suggest that Tic20 and Tic110 have distinct functions. This would explain their parallel existence and their functional relevance (Kovacheva et al. 2005; Hirabayashi et al. 2011).

The TPR domain containing co-chaperone Tic40 recruits stromal chaperones to the translocon (Chou et al. 2003). Tic40 has to be considered as a more recent eukaryotic invention, because it cannot be found in the genomes of sequenced cyanobacteria, red algae, and glaucophytes. The REDOX sensing components Tic32, Tic55, and Tic62 have folds which already existed in cyanobacteria (Balsera et al. 2007; Bodyl et al. 2010; Kalanon and McFadden 2008). The sequences identified in the genomes of glaucophytes and red and green algae, however, are rather distantly related to the sequences found in land plants (Fig. 3). Thus, it is likely that all of these factors are rather late evolutionary achievements.

In conclusion, the evolution of the translocon of the inner envelope has to be envisioned as a bottleneck for establishing the transport of nuclear-coded plastid proteins. SecY remained in the inner envelope for the transport of plastome-encoded proteins like the inner membrane protein Ycf1 (Ladig et al. 2011). LivH is a component of the bacterial leucine/isoleucine transport system (Nazos et al. 1986; Koyanagi et al. 2004), and this makes it likely that it could have participated in the interaction with and may be even in the translocation of unfolded polypeptides. The early occurrence of the translocon component Tic110 suggests that Tic20 was not suitable for the translocation process, at least not for all precursor proteins. The other components, however, evolved to regulate the translocation process with regard to its efficiency (Tic40) and its fine-tuning in response to the metabolic and REDOX conditions of the stroma (Tic32, Tic55, and Tic62).

The Evolution of the Outer Envelope Protein Insertion Machinery

One class of proteins distinct from all those described so far is that of the β-barrel proteins of the outer envelope membrane, which are clearly of prokaryotic origin. The insertion of these proteins appears to be conserved with respect to bacteria by involving Omp85 orthologs (Löffelhardt et al. 2007; Schleiff et al. 2011; Schleiff and Soll 2005). The Omp85 like protein Toc75-V/Oep80 in the plant chloroplast outer envelope (Eckart et al. 2002) is more closely related to the ancestral Omp85 proteins from cyanobacteria than to Toc75-III (Bredemeier et al. 2007) and it was suggested to represent the outer membrane insertase for this pathway (Schleiff et al. 2011; Schleiff and Soll 2005). The gene is essential (Patel et al. 2008; Huang et al. 2011) as expected for a protein involved in the assembly of β-barrel proteins in the outer envelope membrane, but its exact function remains to be determined (Schleiff et al. 2011). The notion that Toc75-V is involved in the assembly of β-barrel proteins is supported by comparison to the mitochondrial system, where the Omp85 homolog Sam50 is involved in β-barrel protein assembly (Kozjak et al. 2003; Paschen et al. 2003; Gentle et al. 2004; Humphries et al. 2005). Sam50 is of clear proteobacterial origin (e.g., Bredemeier et al. 2007). Thus, it

appears very likely that the mechanism for the insertion of β-barrel proteins is conserved between the two organelles and their bacterial relatives, although the recent finding that the POTRA domains of Toc75-V are cytoplasmically exposed argues for the transfer of outer membrane proteins from the cytosol into the membrane (Sommer et al. 2011).

Conclusion

The evolution of a translocation system for the import of plastid proteins was most likely a limiting step after the successful gene transfer from endosymbiont to host genome. Thus, it is not surprising that the translocation path evolved by taking advantage of already existing bacterial proteins that function in the transport of solutes. The outer membrane translocon evolved by inverting and recycling Omp85 to form most likely the ancestral TOC translocon. Tom40, the pore-forming unit of the mitochondrial outer membrane translocon, probably evolved from an ancestral porin-like protein (Gessmann et al. 2011). The difference may be explained by the fact that mitochondria were present at an early time point not requiring a selection of proper targets, while chloroplasts had to discriminate between mitochondrial and chloroplast precursors from the very beginning. The exposure of a domain with chaperone or receptor-like function, which allowed for the recognition of a primitive phenylalanine-based signal, might have been sufficient to discriminate between preproteins targeted towards mitochondria versus chloroplasts. All other components of the chloroplast translocons function only in the translocation event itself. For many factors of the putative initial translocons, bacterial homologs with functions distinct from protein translocation can be identified. Recently, this process was termed "Recycling and Tinkering" (Mirus and Schleiff 2012), which means that existing proteins were placed in a different context and subsequently adapted properties essential for translocation. In the course of evolution, the translocon was shaped to increase specificity and translocation efficiency. The translocon may also have changed to deal with the specific requirements of different plastid types; however, little is known about the role of the translocon in these plastids.

Acknowledgments We are grateful to Stefan Simm, Oliver Mirus, Martina Königer, Andreas Weber, and Arndt von Haeseler for open and critical discussions concerning this issue. Special thanks to Jürgen Steiner, who gave insights into the *Cyanophora paradoxa* genome content prior to publication, which was essential for the description of the evolutionary scenario. The work was supported by grants from the Deutsche Forschungsgemeinschaft SFB807 and from the Volkswagenstiftung to ES.

References

Akita M, Nielsen E, Keegstra K (1997) Identification of protein transport complexes in the chloroplastic envelope membranes via chemical cross-linking. J Cell Biol 136(5):983–994

Allen JF (2003) The function of genomes in bioenergetic organelles. Philos Trans R Soc Lond B Biol Sci 358(1429):19–37, discussion 37–38

Armbrust E, Berges JA, Bowler C, Green BR, Martinez D, Putnam NH, Zhou S, Allen AE, Apt KE, Bechner M, Brzezinski MA, Chaal BK, Chiovitti A, Davis AK, Demarest MS, Detter JC, Glavina T, Goodstein D, Hadi MZ, Hellsten U, Hildebrand M, Jenkins BD, Jurka J, Kapitonov VV, Kroger N, Lau WWY, Lane TW, Larimer FW, Lippmeier JC, Lucas S, Medina M, Montsant A, Obornik M, Parker MS, Palenik B, Pazour GJ, Richardson PM, Rynearson TA, Saito MA, Schwartz DC, Thamatrakoln K, Valentin K, Vardi A, Wilkerson FP, Rokhsar DS (2004) The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. Science 306(5693):79–86

Bae W, Lee Y, Kim DH, Lee J, Kim S, Sohn EJ, Hwang I (2008) AKR2A-mediated import of chloroplast outer membrane proteins is essential for chloroplast biogenesis. Nat Cell Biol 10 (2):220–227

Baldwin AJ, Inoue K (2006) The most C-terminal tri-glycine segment within the polyglycine stretch of the pea Toc75 transit peptide plays a critical role for targeting the protein to the chloroplast outer envelope membrane. FEBS J 273(7):1547–1555

Balsera M, Stengel A, Soll J, Bölter B (2007) Tic62: a protein family from metabolism to protein translocation. BMC Evol Biol 7:43

Beck K, Wu LF, Brunner J, Müller M (2000) Discrimination between SRP- and SecA/SecB-dependent substrates involves selective recognition of nascent chains by SRP and trigger factor. EMBO J 19(1):134–143

Becker T, Jelic M, Vojta A, Radunz A, Soll J, Schleiff E (2004) Preprotein recognition by the Toc complex. EMBO J 23(3):520–530

Becker T, Qbadou S, Jelic M, Schleiff E (2005) Let's talk about . . . chloroplast import. Plant Biol 7 (1):1–14

Benz JP, Stengel A, Lintala M, Lee YH, Weber A, Philippar K, Gugel IL, Kaieda S, Ikegami T, Mulo P, Soll J, Bolter B (2009) *Arabidopsis* Tic62 and ferredoxin-NADP(H) oxidoreductase form light-regulated complexes that are integrated into the chloroplast redox poise. Plant Cell 21(12):3965–3983

Bionda T, Tillmann B, Simm S, Beilstein K, Ruprecht M, Schleiff E (2010) Chloroplast import signals: the length requirement for translocation *in vitro* and *in vivo*. J Mol Biol 402 (3):510–523

Bodyl A, Mackiewicz P, Stiller JW (2010) Comparative genomic studies suggest that the cyanobacterial endosymbionts of the amoeba *Paulinella chromatophora* possess an import apparatus for nuclear-encoded proteins. Plant Biol 12(4):639–649

Bohnsack MT, Schleiff E (2010) The evolution of protein targeting and translocation systems. Biochim Biophys Acta 1803(10):1115–1130

Bonen L, Calixte S (2006) Comparative analysis of bacterial-origin genes for plant mitochondrial ribosomal proteins. Mol Biol Evol 23(3):701–712

Bräutigam A, Weber APM (2009) Proteomic analysis of the proplastid envelope membrane provides novel insights into small molecule and protein transport across proplastid membranes. Mol Plant 2(6):1247–1261

Bredemeier R, Schlegel T, Ertel F, Vojta A, Borissenko L, Bohnsack MT, Groll M, von Haeseler A, Schleiff E (2007) Functional and phylogenetic properties of the pore-forming beta-barrel transporters of the Omp85 family. J Biol Chem 282(3):1882–1890

Bruce BD (2000) Chloroplast transit peptides. Structure, function and evolution. Trends Cell Biol 10(10):440–447

Bruce BD (2001) The paradox of plastid transit peptides: conservation of function despite divergence in primary structure. Biochim Biophys Acta 1541(1–2):2–21

Bullmann L, Haarmann R, Mirus O, Bredemeier R, Hempel F, Maier UG, Schleiff E (2010) Filling the gap, evolutionarily conserved Omp85 in plastids of chromalveolates. J Biol Chem 285 (9):6848–6856

Caliebe A, Grimm R, Kaiser G, Lubeck J, Soll J, Heins L (1997) The chloroplastic protein import machinery contains a Rieske-type iron-sulfur cluster and a mononuclear iron-binding protein. EMBO J 16(24):7342–7350

Cavalier-Smith T (2003) Genomic reduction and evolution of novel genetic membranes and protein-targeting machinery in eukaryote-eukaryote chimaeras (meta-algae). Philos Trans R Soc Lond B Biol Sci 358(1429):109–133, discussion 133–134

Chew O, Lister R, Qbadou S, Heazlewood JL, Soll J, Schleiff E, Millar AH, Whelan J (2004) A plant outer mitochondrial membrane protein with high amino acid sequence identity to a chloroplast protein import receptor. FEBS Lett 557(1–3):109–114

Chiu CC, Li HM (2008) Tic40 is important for reinsertion of proteins from the chloroplast stroma into the inner membrane. Plant J 56(5):793–801

Chou ML, Fitzpatrick LM, Tu SL, Budziszewski G, Potter-Lewis S, Akita M, Levin JZ, Keegstra K, Li HM (2003) Tic40, a membrane-anchored co-chaperone homolog in the chloroplast protein translocon. EMBO J 22(12):2970–2980

Cline K, Dabney-Smith C (2008) Plastid protein import and sorting: different paths to the same compartments. Curr Opin Plant Biol 11(6):585–592

Deusch O, Landan G, Roettger M, Gruenheit N, Kowallik KV, Allen JF, Martin W, Dagan T (2008) Genes of cyanobacterial origin in plant nuclear genomes point to a heterocyst-forming plastid ancestor. Mol Biol Evol 25(4):748–761

Driessen AJM, Nouwen N (2008) Protein translocation across the bacterial cytoplasmic membrane. Annu Rev Biochem 77:643–667

Driessen AJ, Manting EH, van der Does C (2001) The structural basis of protein targeting and translocation in bacteria. Nat Struct Biol 8(6):492–498

du Plessis DJF, Nouwen N, Driessen AJM (2011) The Sec translocase. Biochim Biophys Acta 1808(3):851–865

Duy D, Wanner G, Meda AR, von Wiren N, Soll J, Philippar K (2007) PIC1, an ancient permease in *Arabidopsis* chloroplasts, mediates iron transport. Plant Cell 19(3):986–1006

Eckart K, Eichacker L, Sohrt K, Schleiff E, Heins L, Soll J (2002) A Toc75-like protein import channel is abundant in chloroplasts. EMBO Rep 3(6):557–562

Emanuelsson O, Nielsen H, Brunak S, von Heijne G (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J Mol Biol 300(4):1005–1016

Ertel F, Mirus O, Bredemeier R, Moslavac S, Becker T, Schleiff E (2005) The evolutionarily related beta-barrel polypeptide transporters from *Pisum sativum* and *Nostoc* PCC7120 contain two distinct functional domains. J Biol Chem 280(31):28281–28289

Fellerer C, Schweiger R, Schongruber K, Soll J, Schwenkert S (2011) Cytosolic HSP90 cochaperones HOP and FKBP interact with freshly synthesized chloroplast preproteins of *Arabidopsis*. Mol Plant 4(6):1133–1145

Field MC, Dacks JB (2009) First and last ancestors: reconstructing evolution of the endomembrane system with ESCRTs, vesicle coat proteins, and nuclear pore complexes. Curr Opin Cell Biol 21(1):4–13

Fulda S, Norling B, Schoor A, Hagemann M (2002) The Slr0924 protein of *Synechocystis* sp. strain PCC 6803 resembles a subunit of the chloroplast protein import complex and is mainly localized in the thylakoid lumen. Plant Mol Biol 49(1):107–118

Gentle I, Gabriel K, Beech P, Waller R, Lithgow T (2004) The Omp85 family of proteins is essential for outer membrane biogenesis in mitochondria and bacteria. J Cell Biol 164 (1):19–24

Gessmann D, Flinner N, Pfannstiel J, Schlosinger A, Schleiff E, Nussberger S, Mirus O (2011) Structural elements of the mitochondrial preprotein-conducting channel Tom40 dissolved by bioinformatics and mass spectrometry. Biochim Biophys Acta 1807(12):1647–1657

Gould SB, Sommer MS, Hadfi K, Zauner S, Kroth PG, Maier UG (2006) Protein targeting into the complex plastid of cryptophytes. J Mol Evol 62(6):674–681

Grosche C, Hempel F, Bolte K, Abram L, Maier UG, Zauner S (2013) Protein import into complex plastids: current findings and perspectives. In: Löffelhardt W (ed) Endosymbiosis. Springer, Wien New York, pp 215–232

Grudnik P, Bange G, Sinning I (2009) Protein targeting by the signal recognition particle. Biol Chem 390(8):775–782

Gupta RS (1995) Phylogenetic analysis of the 90 kD heat shock family of protein sequences and an examination of the relationship among animals, plants, and fungi species. Mol Biol Evol 12 (6):1063–1073

Gupta RS, Golding GB (1993) Evolution of HSP70 gene and its implications regarding relationships between archaebacteria, eubacteria, and eukaryotes. J Mol Evol 37(6):573–582

Hagan CL, Silhavy TJ, Kahne D (2011) beta-Barrel membrane protein assembly by the Bam complex. Annu Rev Biochem 80:189–210

Harb OS, Chatterjee B, Fraunholz MJ, Crawford MJ, Nishi M, Roos DS (2004) Multiple functionally redundant signals mediate targeting to the apicoplast in the apicomplexan parasite *Toxoplasma gondii*. Eukaryot Cell 3(3):663–674

Heins L, Mehrle A, Hemmler R, Wagner R, Kuchler M, Hormann F, Sveshnikov D, Soll J (2002) The preprotein conducting channel at the inner envelope membrane of plastids. EMBO J 21 (11):2616–2625

Hinnah SC, Wagner R, Sveshnikova N, Harrer R, Soll J (2002) The chloroplast protein import channel Toc75: pore properties and interaction with transit peptides. Biophys J 83(2):899–911

Hirabayashi Y, Kikuchi S, Oishi M, Nakai M (2011) *In vivo* studies on the roles of two closely related *Arabidopsis* Tic20 proteins, AtTic20-I and AtTic20-IV. Plant Cell Physiol 52 (3):469–478

Hirohashi T, Hase T, Nakai M (2001) Maize non-photosynthetic ferredoxin precursor is mis-sorted to the intermembrane space of chloroplasts in the presence of light. Plant Physiol 125 (4):2154–2163

Hiss JA, Schneider G (2009) Architecture, function and prediction of long signal peptides. Brief Bioinform 10(5):569–578

Huang W, Ling Q, Bedard J, Lilley K, Jarvis P (2011) In vivo analyses of the roles of essential Omp85-related proteins in the chloroplast outer envelope membrane. Plant Physiol 157 (1):147–159

Humphries AD, Streimann IC, Stojanovski D, Johnston AJ, Yano M, Hoogenraad NJ, Ryan MT (2005) Dissection of the mitochondrial import and assembly pathway for human Tom40. J Biol Chem 280(12):11535–11543

Ieva R, Tian P, Peterson JH, Bernstein HD (2011) Sequential and spatially restricted interactions of assembly factors with an autotransporter beta domain. Proc Natl Acad Sci USA 108(31): E383–E391

Inoue K, Keegstra K (2003) A polyglycine stretch is necessary for proper targeting of the protein translocation channel precursor to the outer envelope membrane of chloroplasts. Plant J 34 (5):661–669

Inoue K, Potter D (2004) The chloroplastic protein translocation channel Toc75 and its paralog OEP80 represent two distinct protein families and are targeted to the chloroplastic outer envelope by different mechanisms. Plant J 39(3):354–365

Inoue K, Demel R, de Kruijff B, Keegstra K (2001) The N-terminal portion of the preToc75 transit peptide interacts with membrane lipids and inhibits binding and import of precursor proteins into isolated chloroplasts. Eur J Biochem 268(14):4036–4043

Inoue K, Baldwin AJ, Shipman RL, Matsui K, Theg SM, Ohme-Takagi M (2005) Complete maturation of the plastid protein translocation channel requires a type I signal peptidase. J Cell Biol 171(3):425–430

Jackson-Constan D, Akita M, Keegstra K (2001) Molecular chaperones involved in chloroplast protein import. Biochim Biophys Acta 1541(1–2):102–113

Jacob-Dubuisson F, Villeret V, Clantin B, Delattre AS, Saint N (2009) First structural insights into the TpsB/Omp85 superfamily. Biol Chem 390(8):675–684

Jarvis P (2008) Targeting of nucleus-encoded proteins to chloroplasts in plants. New Phytol 179 (2):257–285

Jürgens G (2004) Membrane trafficking in plants. Annu Rev Cell Dev Biol 20:481–504

Kalanon M, McFadden GI (2008) The chloroplast protein translocation complexes of *Chlamydomonas reinhardtii*. a bioinformatic comparison of Toc and Tic components in plants, green algae and red algae. Genetics 179(1):95–112

Kessler F, Schnell D (2009) Chloroplast biogenesis: diversity and regulation of the protein import apparatus. Curr Opin Cell Biol 21(4):494–500

Kikuchi S, Oishi M, Hirabayashi Y, Lee DW, Hwang I, Nakai M (2009) A 1-megadalton translocation complex containing Tic20 and Tic21 mediates chloroplast protein import at the inner envelope membrane. Plant Cell 21(6):1781–1797

Kikuchi S, Bédard J, Hirano M, Hirabayashi Y, Oishi M, Imai M, Takase M, Ide T, Nakai M (2013) Uncovering the protein translocon at the chloroplast inner envelope membrane. Science 339(6119):571–574

Kilian O, Kroth PG (2005) Identification and characterization of a new conserved motif within the presequence of proteins targeted into complex diatom plastids. Plant J 41(2):175–183

Kleffmann T, Russenberger D, von Zychlinski A, Christopher W, Sjolander K, Gruissem W, Baginsky S (2004) The *Arabidopsis thaliana* chloroplast proteome reveals pathway abundance and novel protein functions. Curr Biol 14(5):354–362

Knowles TJ, Scott-Tucker A, Overduin M, Henderson IR (2009) Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. Nat Rev Microbiol 7 (3):206–214

Koenig P, Mirus O, Haarmann R, Sommer MS, Sinning I, Schleiff E, Tews I (2010) Conserved properties of polypeptide transport-associated (POTRA) domains derived from cyanobacterial Omp85. J Biol Chem 285(23):18016–18024

Kouranov A, Chen X, Fuks B, Schnell DJ (1998) Tic20 and Tic22 are new components of the protein import apparatus at the chloroplast inner envelope membrane. J Cell Biol 143 (4):991–1002

Kovacheva S, Bedard J, Patel R, Dudley P, Twell D, Rios G, Koncz C, Jarvis P (2005) In vivo studies on the roles of Tic110, Tic40 and Hsp93 during chloroplast protein import. Plant J 41 (3):412–428

Kovacheva S, Bedard J, Wardle A, Patel R, Jarvis P (2007) Further in vivo studies on the role of the molecular chaperone, Hsp93, in plastid protein import. Plant J 50(2):364–379

Kovacs-Bogdan E, Benz JP, Soll J, Bölter B (2011) Tic20 forms a channel independent of Tic110 in chloroplasts. BMC Plant Biol 11:133

Koyanagi T, Katayama T, Suzuki H, Kumagai H (2004) Identification of the LIV-I/LS system as the third phenylalanine transporter in *Escherichia coli* K-12. J Bacteriol 186(2):343–350

Kozjak V, Wiedemann N, Milenkovic D, Lohaus C, Meyer HE, Guiard B, Meisinger C, Pfanner N (2003) An essential role of Sam50 in the protein sorting and assembly machinery of the mitochondrial outer membrane. J Biol Chem 278(49):48520–48523

Kuchler M, Decker S, Hormann F, Soll J, Heins L (2002) Protein import into chloroplasts involves redox-regulated proteins. EMBO J 21(22):6136–6145

Ladig R, Sommer MS, Hahn A, Leisegang MS, Papasotiriou DG, Ibrahim M, Elkehal R, Karas M, Zickermann V, Gutensohn M, Brandt U, Klosgen RB, Schleiff E (2011) A high-definition native polyacrylamide gel electrophoresis system for the analysis of membrane complexes. Plant J 67(1):181–194

Lee DW, Lee S, Lee GJ, Lee KH, Kim S, Cheong GW, Hwang I (2006) Functional characterization of sequence motifs in the transit peptide of *Arabidopsis* small subunit of Rubisco. Plant Physiol 140(2):466–483

Lee DW, Kim JK, Lee S, Choi S, Kim S, Hwang I (2008) *Arabidopsis* nuclear-encoded plastid transit peptides contain multiple sequence subgroups with distinctive chloroplast-targeting sequence motifs. Plant Cell 20(6):1603–1622

Lee S, Lee DW, Lee Y, Mayer U, Stierhof YD, Lee S, Jürgens G, Hwang I (2009) Heat shock protein cognate 70-4 and an E3 ubiquitin ligase, CHIP, mediate plastid-destined precursor degradation through the ubiquitin-26S proteasome system in *Arabidopsis*. Plant Cell 21 (12):3984–4001

Leipe DD, Wolf YI, Koonin EV, Aravind L (2002) Classification and evolution of P-loop GTPases and related ATPases. J Mol Biol 317(1):41–72

Li HM, Chiu CC (2010) Protein transport into chloroplasts. Annu Rev Plant Biol 61:157–180

Li M, Schnell DJ (2006) Reconstitution of protein targeting to the inner envelope membrane of chloroplasts. J Cell Biol 175(2):249–259

Löffelhardt W, von Haeseler A, Schleiff E (2007) The beta-barrel shaped polypeptide transporter, an old concept for precursor protein transfer across membranes. Symbiosis 44:33–42

Lubeck J, Soll J, Akita M, Nielsen E, Keegstra K (1996) Topology of IEP110, a component of the chloroplastic protein import machinery present in the inner envelope membrane. EMBO J 15 (16):4230–4238

Lubeck J, Heins L, Soll J (1997) A nuclear-coded chloroplastic inner envelope membrane protein uses a soluble sorting intermediate upon import into the organelle. J Cell Biol 137 (6):1279–1286

Lv HX, Guo GQ, Yang ZN (2009) Translocons on the inner and outer envelopes of chloroplasts share similar evolutionary origin in *Arabidopsis thaliana*. J Evol Biol 22(7):1418–1428

Mandon EC, Trueman SF, Gilmore R (2009) Translocation of proteins through the Sec61 and SecYEG channels. Curr Opin Cell Biol 21(4):501–507

Marshall JS, DeRocher AE, Keegstra K, Vierling E (1990) Identification of heat shock protein hsp70 homologues in chloroplasts. Proc Natl Acad Sci USA 87(1):374–378

Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, Leister D, Stoebe B, Hasegawa M, Penny D (2002) Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proc Natl Acad Sci USA 99(19):12246–12251

Mergaert P, Nikovics K, Kelemen Z, Maunoury N, Vaubert D, Kondorosi A, Kondorosi E (2003) A novel family in *Medicago truncatula* consisting of more than 300 nodule-specific genes coding for small, secreted polypeptides with conserved cysteine motifs. Plant Physiol 132 (1):161–173

Miras S, Salvi D, Ferro M, Grunwald D, Garin J, Joyard J, Rolland N (2002) Non-canonical transit peptide for import into the chloroplast. J Biol Chem 277(49):47770–47778

Miras S, Salvi D, Piette L, Seigneurin-Berny D, Grunwald D, Reinbothe C, Joyard J, Reinbothe S, Rolland N (2007) Toc159- and Toc75-independent import of a transit sequence-less precursor into the inner envelope of chloroplasts. J Biol Chem 282(40):29482–29492

Mirus O, Schleiff E (2009) The evolution of tetratricopeptide repeat domain containing receptors involved in protein translocation. Endocytobiosis Cell Res 19:31–50

Mirus O, Schleiff E (2012) Recycling and tinkering - the evolution of protein transport to and into endosymbiotically derived organelles. In: Bullerwell CE (ed) Organelle genetics - evolution of organelle genomes and gene expression. Springer, Berlin

Mirus O, Bionda T, von Haeseler A, Schleiff E (2009) Evolutionarily evolved discriminators in the 3-TPR domain of the Toc64 family involved in protein translocation at the outer membrane of chloroplasts and mitochondria. J Mol Model 15(8):971–982

Mirus O, Hahn A, Schleiff E (2010) Outer membrane proteins. In: König H, Claus H, Varma A (eds) Prokaryotic cell wall compounds. Structure and biochemistry. Springer, Berlin, pp 175–231

Moslavac S, Mirus O, Bredemeier R, Soll J, von Haeseler A, Schleiff E (2005) Conserved pore-forming regions in polypeptide-transporting proteins. FEBS J 272(6):1367–1378

Natale P, Bruser T, Driessen AJM (2008) Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane - distinct translocases and mechanisms. Biochim Biophys Acta 1778(9):1735–1756

Nazos PM, Antonucci TK, Landick R, Oxender DL (1986) Cloning and characterization of *livH*, the structural gene encoding a component of the leucine transport system in *Escherichia coli*. J Bacteriol 166(2):565–573

Nicolaisen K, Mariscal V, Bredemeier R, Pernil R, Moslavac S, Lopez-Igual R, Maldener I, Herrero A, Schleiff E, Flores E (2009) The outer membrane of a heterocyst-forming cyanobacterium is a permeability barrier for uptake of metabolites that are exchanged between cells. Mol Microbiol 74(1):58–70

Nielsen E, Akita M, Davila-Aponte J, Keegstra K (1997) Stable association of chloroplastic precursors with protein translocation complexes that contain proteins from both envelope membranes and a stromal Hsp100 molecular chaperone. EMBO J 16(5):935–946

Nowack EC, Grossman AR (2012) Trafficking of protein into the recently established photosynthetic organelles of *Paulinella chromatophora*. Proc Natl Acad Sci USA 109:5340–5345

Oreb M, Reger K, Schleiff E (2006) Chloroplast protein import. Reverse genetic approaches. Curr Genomics 7:235–244

Oreb M, Tews I, Schleiff E (2008) Policing Tic 'n' Toc, the doorway to chloroplasts. Trends Cell Biol 18(1):19–27

Palmer T, Sargent F, Berks BC (2005) Export of complex cofactor-containing proteins by the bacterial Tat pathway. Trends Microbiol 13(4):175–180

Paschen SA, Waizenegger T, Stan T, Preuss M, Cyrklaff M, Hell K, Rapaport D, Neupert W (2003) Evolutionary conservation of biogenesis of beta-barrel membrane proteins. Nature 426 (6968):862–866

Patel R, Hsu SC, Bedard J, Inoue K, Jarvis P (2008) The Omp85-related chloroplast outer envelope protein OEP80 is essential for viability in Arabidopsis. Plant Physiol 148(1):235–245

Patron NJ, Waller RF (2007) Transit peptide diversity and divergence: a global analysis of plastid targeting signals. Bioessays 29(10):1048–1058

Petersen TN, Brunak S, von Heijne G, Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 8(10):785–786

Pohlschroder M, Hartmann E, Hand NJ, Dilks K, Haddad A (2005) Diversity and evolution of protein translocation. Annu Rev Microbiol 59:91–111

Price DC, Chan CX, Yoon HS, Yang E, Qiu H, Weber AP, Schwacke R, Gross J, Blouin N, Lane C, Reyes-Prieto A, Durnford D, Neilson J, Lang B, Burger G, Steiner J, Loffelhardt W, Meuser J, Posewitz M, Ball S, Arias M, Henrissat B, Coutinho P, Rensing S, Symeonidi A, Doddapaneni H, Green B, Rajah V, Boore J, Bhattacharya D (2012) *Cyanophora paradoxa* genome elucidates origin of photosynthesis in algae and plants. Science 335(6070):843–847

Qbadou S, Becker T, Mirus O, Tews I, Soll J, Schleiff E (2006) The molecular chaperone Hsp90 delivers precursor proteins to the chloroplast import receptor Toc64. EMBO J 25 (9):1836–1847

Qbadou S, Becker T, Bionda T, Reger K, Ruprecht M, Soll J, Schleiff E (2007) Toc64–a preprotein-receptor at the outer membrane with bipartide function. J Mol Biol 367 (5):1330–1346

Ralph SA, Foth BJ, Hall N, McFadden GI (2004) Evolutionary pressures on apicoplast transit peptides. Mol Biol Evol 21(12):2183–2194

Reinbothe S, Quigley F, Gray J, Schemenewitz A, Reinbothe C (2004) Identification of plastid envelope proteins required for import of protochlorophyllide oxidoreductase A into the chloroplast of barley. Proc Natl Acad Sci USA 101(7):2197–2202

Richardson LGI, Jelokhani-Niaraki M, Smith MD (2009) The acidic domains of the Toc159 chloroplast preprotein receptor family are intrinsically disordered protein domains. BMC Biochem 10:35

Robinson C, Bolhuis A (2001) Protein targeting by the twin-arginine translocation pathway. Nat Rev Mol Cell Biol 2(5):350–356

Rolland N, Job D, Douce R (1993) Common sequence motifs coding for higher-plant and prokaryotic O-acetylserine (thiol)-lyases: bacterial origin of a chloroplast transit peptide? Biochem J 293(Pt 3):829–833

Ruiz N, Kahne D, Silhavy T (2006) Advances in understanding bacterial outer-membrane biogenesis. Nat Rev Microbiol 4(1):57–66

Ruprecht M, Bionda T, Sato T, Sommer MS, Endo T, Schleiff E (2010) On the impact of precursor unfolding during protein import into chloroplasts. Mol Plant 3(3):499–508

Rusch SL, Kendall DA (2007) Interactions that drive Sec-dependent bacterial protein transport. Biochemistry 46(34):9665–9673

Scheufler C, Brinker A, Bourenkov G, Pegoraro S, Moroder L, Bartunik H, Hartl FU, Moarefi I (2000) Structure of TPR domain-peptide complexes. Critical elements in the assembly of the Hsp70-Hsp90 multichaperone machine. Cell 101(2):199–210

Schlegel T, Mirus O, von Haeseler A, Schleiff E (2007) The tetratricopeptide repeats of receptors involved in protein translocation across membranes. Mol Biol Evol 24(12):2763–2774

Schleiff E, Becker T (2011) Common ground for protein translocation. Access control for mitochondria and chloroplasts. Nat Rev Mol Cell Biol 12(1):48–59

Schleiff E, Soll J (2000) Travelling of proteins through membranes: translocation into chloroplasts. Planta 211(4):449–456

Schleiff E, Soll J (2005) Membrane protein insertion: mixing eukaryotic and prokaryotic concepts. EMBO Rep 6(11):1023–1027

Schleiff E, Eichacker LA, Eckart K, Becker T, Mirus O, Stahl T, Soll J (2003) Prediction of the plant beta-barrel proteome. A case study of the chloroplast outer envelope. Protein Sci 12 (4):748–759

Schleiff E, Maier UG, Becker T (2011) Omp85 in eukaryotic systems: one protein family with distinct functions. Biol Chem 392(1–2):21–27

Schnell DJ, Kessler F, Blobel G (1994) Isolation of components of the chloroplast protein import machinery. Science 266(5187):1007–1012

Schwenkert S, Soll J, Bölter B (2011) Protein import into chloroplasts–how chaperones feature into the game. Biochim Biophys Acta 1808(3):901–911

Scott SV, Theg SM (1996) A new chloroplast protein import intermediate reveals distinct translocation machineries in the two envelope membranes: energetics and mechanistic implications. J Cell Biol 132(1–2):63–75

Shi L, Theg S (2010) A stromal heat shock protein 70 system functions in protein import into chloroplasts in the moss *Physcomitrella patens*. Plant Cell 22(1):205–220

Skalitzky CA, Martin JR, Harwood JH, Beirne JJ, Adamczyk BJ, Heck GR, Cline K, Fernandez DE (2011) Plastids contain a second sec translocase system with essential functions. Plant Physiol 155(1):354–369

Sklar JG, Wu T, Kahne D, Silhavy TJ (2007) Defining the roles of the periplasmic chaperones SurA, Skp, and DegP in Escherichia coli. Genes Dev 21(19):2473–2484

Sohrt K, Soll J (2000) Toc64, a new component of the protein translocon of chloroplasts. J Cell Biol 148(6):1213–1221

Soll J, Schleiff E (2004) Protein import into chloroplasts. Nat Rev Mol Cell Biol 5(3):198–208

Sommer MS, Schleiff E (2009) Molecular interactions within the plant TOC complex. Biol Chem 390(8):739–744

Sommer MS, Daum B, Gross LE, Weis BL, Mirus O, Abram L, Maier UG, Kühlbrandt W, Schleiff E (2011) Chloroplast Omp85 proteins change orientation during evolution. Proc Natl Acad Sci USA 108(33):13841–13846

Spence E, Sarcina M, Ray N, Møller SG, Mullineaux CW, Robinson C (2003) Membrane-specific targeting of green fluorescent protein by the Tat pathway in the cyanobacterium *Synechocystis* PCC6803. Mol Microbiol 48(6):1481–1489

Stahl T, Glockmann C, Soll J, Heins L (1999) Tic40, a new "old" subunit of the chloroplast protein import translocon. J Biol Chem 274(52):37467–37472

Stegemann S, Bock R (2006) Experimental reconstruction of functional gene transfer from the tobacco plastid genome to the nucleus. Plant Cell 18(11):2869–2878
Stegemann S, Hartmann S, Ruf S, Bock R (2003) High-frequency gene transfer from the chloroplast genome to the nucleus. Proc Natl Acad Sci USA 100(15):8828–8833
Steiner JM, Löffelhardt W (2005) Protein translocation into and within cyanelles (Review). Mol Membr Biol 22(1–2):123–132
Steiner JM, Yusa F, Pompe JA, Löffelhardt W (2005) Homologous protein import machineries in chloroplasts and cyanelles. Plant J 44(4):646–652
Stirling PC, Bakhoum SF, Feigl AB, Leroux MR (2006) Convergent evolution of clamp-like binding sites in diverse chaperones. Nat Struct Mol Biol 13(10):865–870
Struyve M, Moons M, Tommassen J (1991) Carboxy-terminal phenylalanine is essential for the correct assembly of a bacterial outer membrane protein. J Mol Biol 218(1):141–148
Stymest KH, Klappa P (2008) The periplasmic peptidyl prolyl cis-trans isomerases PpiD and SurA have partially overlapping substrate specificities. FEBS J 275(13):3470–3479
Su PH, Li HM (2010) Stromal Hsp70 is important for protein translocation into pea and Arabidopsis chloroplasts. Plant Cell 22(5):1516–1531
Teng Y, Su Y, Chen L, Lee Y, Hwang I, Li H (2006) Tic21 is an essential translocon component for protein translocation across the chloroplast inner envelope membrane. Plant Cell 18 (9):2247–2257
Tranel PJ, Keegstra K (1996) A novel, bipartite transit peptide targets OEP75 to the outer membrane of the chloroplastic envelope. Plant Cell 8(11):2093–2104
Tranel PJ, Froehlich J, Goyal A, Keegstra K (1995) A component of the chloroplastic protein import apparatus is targeted to the outer envelope membrane via a novel pathway. EMBO J 14 (11):2436–2446
Tripp J, Hahn A, Koenig P, Flinner N, Bublak D, Brouwer EM, Ertel F, Mirus O, Sinning I, Tews I, Schleiff E (2012) Structure and conservation of the periplasmic targeting factor Tic22 protein from plants and cyanobacteria. J Biol Chem 287(29):24164–24173
Tsukazaki T, Mori H, Fukai S, Ishitani R, Mori T, Dohmae N, Perederina A, Sugita Y, Vassylyev DG, Ito K, Nureki O (2008) Conformational transition of Sec machinery inferred from bacterial SecYE structures. Nature 455(7215):988–991
Ullers RS, Luirink J, Harms N, Schwager F, Georgopoulos C, Genevaux P (2004) SecB is a bona fide generalized chaperone in *Escherichia coli*. Proc Natl Acad Sci USA 101(20):7583–7588
van de Velde W, Zehirov G, Szatmari A, Debreczeny M, Ishihara H, Kevei Z, Farkas A, Mikulass K, Nagy A, Tiricz H, Satiat-Jeunemaitre B, Alunni B, Bourge M, Kucho KI, Abe M, Kereszt A, Maroti G, Uchiumi T, Kondorosi E, Mergaert P (2010) Plant peptides govern terminal differentiation of bacteria in symbiosis. Science 327(5969):1122–1126
van der Giezen M, Tovar J (2005) Degenerate mitochondria. EMBO Rep 6(6):525–530
van der Laan M, Nouwen NP, Driessen AJM (2005) YidC - an evolutionary conserved device for the assembly of energy-transducing membrane protein complexes. Curr Opin Microbiol 8 (2):182–187
van Dooren GG, Schwartzbach SD, Osafune T, McFadden GI (2001) Translocation of proteins across the multiple membranes of complex plastids. Biochim Biophys Acta 1541(1–2):34–53
Verma DP, Hong Z (1996) Biogenesis of the peribacteroid membrane in root nodules. Trends Microbiol 4(9):364–368
Villarejo A, Buren S, Larsson S, Dejardin A, Monne M, Rudhe C, Karlsson J, Jansson S, Lerouge P, Rolland N, von Heijne G, Grebe M, Bako L, Samuelsson G (2005) Evidence for a protein transported through the secretory pathway en route to the higher plant chloroplast. Nat Cell Biol 7(12):1224–1231
Vothknecht UC, Soll J (2005) Chloroplast membrane transport: interplay of prokaryotic and eukaryotic traits. Gene 354:99–109
Wolf DH, Stolz A (2012) The Cdc48 machine in endoplasmic reticulum associated protein degradation. Biochim Biophys Acta 1823(1):117–124

Wunder T, Martin R, Löffelhardt W, Schleiff E, Steiner JM (2007) The invariant phenylalanine of precursor proteins discloses the importance of Omp85 for protein translocation into cyanelles. BMC Evol Biol 7:236

Yusa F, Steiner JM, Löffelhardt W (2008) Evolutionary conservation of dual Sec translocases in the cyanelles of *Cyanophora paradoxa*. BMC Evol Biol 8:304

Zak E, Norling B, Maitra R, Huang F, Andersson B, Pakrasi HB (2001) The initial steps of biogenesis of cyanobacterial photosystems occur in plasma membranes. Proc Natl Acad Sci USA 98(23):13443–13448

Zhang XP, Glaser E (2002) Interaction of plant mitochondrial and chloroplast signal peptides with the Hsp70 molecular chaperone. Trends Plant Sci 7(1):14–21

Evolution of Storage Polysaccharide Metabolism in Archaeplastida Opens an Unexpected Window on the Molecular Mechanisms That Drove Plastid Endosymbiosis

S.G. Ball

Abstract Plastid endosymbiosis was selected through the establishment of a biochemical link between the disconnected metabolic networks of the cyanobiont and its eukaryote host. This link is likely to have consisted of the efflux of photosynthetic carbon from the bacterial symbiont to the cytosol of the eukaryote. Storage molecules are suspected to have played a pivotal role in the establishment of such a flux. The latter provided an immediate opportunity to feed carbon upon a supply dictated by cyanobacterial metabolism while allowing the host to tap these resources upon demand according to its own regulatory circuits. The presence of the stores thus buffered the disconnected and unrelated sources and sink pathways of photosynthetic carbon metabolism during the early stages of plastid endosymbiosis. Comparisons of extant biochemical networks explaining storage polysaccharide metabolism in the three lineages that emerged after plastid endosymbiosis have enabled the reconstruction of the simplest hypothetical ancient network. The latter possibly consisted of the export of photosynthate from the cyanobiont in the form of the bacteria-specific metabolite ADP-glucose and the polymerization of the latter in the host cytosol through an ADP-glucose-specific glucan synthase. Neither the required ADP-glucose transporter nor the glucan synthase can be suspected to have been encoded by the cyanobacterial or host genomes prior to endosymbiosis. Nevertheless these critical components were required to trigger the event. The possible origin of these two key proteins is reviewed.

S.G. Ball (✉)
Unité de Glycobiologie Structurale et Fonctionnelle, Université des Sciences et Technologies de Lille, UMR8576 du CNRS, Lille, France
e-mail: Steven.Ball@univ-lille1.fr

W. Löffelhardt (ed.), *Endosymbiosis*, DOI 10.1007/978-3-7091-1303-5_6,

© Springer-Verlag Wien 2014

Contents

Introduction 112
Storage: The Metabolic Answer to Connect Unrelated Biochemical Networks 113
The Comparative Biochemistry of Starch and Glycogen Metabolism 113
Starch Metabolism in Archaeplastida 116
Reconstruction of the Ancient Network of Starch Metabolism 121
Incompatibility of Phototrophy and Diazotrophy in Photosynthetic Eukaryotes 124
The ADP-Glucose Connection 127
ADP-Glucose Transport and Glucan Polymerization 129
References 131

Introduction

Oxygenic photosynthesis is a remarkably complex mechanism that evolved in cyanobacteria and durably changed the chemistry of the atmosphere through the massive release of oxygen from water. This very efficient mechanism of tapping energy from the sun to reduce carbon dioxide to organic components requires a vast number of proteins interacting in large complexes finely tuned to avoid the production of toxic waste. The complexity of the process made the progressive introgression to non-cyanobacterial taxa of the hundreds of genes required to achieve it an impossible task. Yet oxygenic photosynthesis was transmitted to eukaryotes allowing primary producers to benefit from the infinite possibilities offered either by eukaryotic multicellularity programs or mobility or by the ability of some of these organisms to colonize environments inaccessible to cyanobacteria. Later on, eukaryotic photosynthesizers were able to maximize primary production by the building of tall rigid structures, thus paving the way for the bloom of animal life. But how did eukaryotes succeed where non-cyanobacterial prokaryotes failed? How did they capture the hundreds of required cyanobacterial genes? The evolution of phagotrophy in early eukaryotes offered a unique opportunity for plant ancestors to capture cyanobacteria as prey (Raven et al. 2009). It is plausible that aborted phagotrophic events enabled stable symbiotic interactions between prey and predator. A particular eukaryotic phagotroph eventually established a specific cyanobacterial lineage as endosymbiont and went down the path of metabolic integration of the cyanobiont into a true cellular organelle: the plastid (for a general review, see Chan et al. 2011). This review focuses on those molecular events that established the initial carbon flux of photosynthesis at endosymbiosis.

Storage: The Metabolic Answer to Connect Unrelated Biochemical Networks

If we hypothesize that plastid endosymbiosis resulted from an aborted phagocytosis, then there is no reason to suppose the preexistence of any molecular connection between the ingested prey and its predator. If we further assume that plastid endosymbiosis was selected because it provided photosynthate to heterotrophic eukaryotes, then we can appreciate the magnitude of the problems facing the partners of plastid endosymbiosis. An optimized metabolic flux exporting carbon from the cyanobiont to its host had to be set up at the very beginning of the process. There was very little time for innovations and no time for adaptation to the problem of unsynchronized demand and supply of carbon. The host had no way of signaling the cyanobiont that it needed carbon nor could it be signaled that carbon could be made accessible by the cyanobiont. Yet the flux came directly under fierce natural selection and was responsible for the initial success of the organisms that had achieved endosymbiosis. How could this be done? It is unlikely that tapping just any metabolite out of the cyanobiont would be without consequences on biochemical networks that have undergone millions year of selection to work in an optimized fashion. Moreover it is equally unlikely that a sudden burst of metabolite in the cytosol would be neutral with respect to the highly coordinated cytosolic networks. An obvious solution to this problem would be to store the carbon in a form readily accessible to the host through catabolic networks defined by preexisting and thus optimized host catabolism responding solely to host needs. Because the size of the stored carbon pools was likely to vary widely as a function of unsynchronized carbon demand and supply the need emerged for a pool having little if any impact on host osmolarity. Storage polysaccharide metabolism therefore defines an obvious candidate for the establishment of the first connection between the partners of endosymbiosis and examining the evolution of this pathway in Archaeplastida and their cyanobacterial and eukaryote ancestors may provide useful information in this respect. In this review we will propose that plastid endosymbiosis has indeed been established through the forging of a metabolic link making uses of the similarities and differences between the preexisting networks of host and cyanobiont storage polysaccharide metabolism. To fully grasp the issues we will briefly first outline what these similarities and differences were.

The Comparative Biochemistry of Starch and Glycogen Metabolism

Both cyanobacterial and host partners can be safely assumed to have been alpha-glucan accumulators before endosymbosis (Ball et al. 2011). This can be deduced from the finding of enzymes of starch metabolism in all Archaeplastida that are both of distinctive cyanobacterial phylogenetic origin and of eukaryotic host origin

(Coppin et al. 2005; Patron and Keeling 2005; Deschamps et al. 2008a). In addition all extant free-living cyanobacteria are either glycogen or starch accumulators (Nakamura et al. 2005) and many heterotrophic eukaryote lineages accumulate glycogen (Ball et al. 2011).

Storage α-glucans come in a variety of forms including maltooligosaccharides (MOS), glycogen, and starch (Boos and Shuman 1998; Buléon et al. 1998). Glycogen defines by far the most widespread form of storage polysaccharide found in all three domains of life (eukaryotes archaea and bacteria). Glycogen consists of small hydrosoluble particles with a maximal diameter of 40–50 nm made of glucose linked by α-1,4 glycosidic linkages with 8–12 % α-1,6 branches. Each branch creates a novel chain increasing progressively the density of the glucans (glucose chains) at the periphery of the particle and therefore limiting its size to less than 50 nm in diameter. The glycogen outer chains are readily accessible to enzymes of glycogen synthesis and degradation as if the glucose was in a soluble state. Hence glycogen defines a very dynamic form of storing glucose yet with little osmotic activity while remaining rapidly available to cellular metabolism (for a review of glycogen structure, see Buléon et al. 1998; Shearer and Graham 2002). Starch has the same composition and chemical linkages but a very different physicochemical state. It defines a solid insoluble semicrystalline structure inert osmotically but also unavailable to hydrosoluble enzymes. Polysaccharide crystallization is ensured by an ordered distribution of branches which remain concentrated in certain regions of the polysaccharide (for a detailed review of starch structure, see Buléon et al. 1998). This asymmetric distribution leads to a greater proximity of chains into clusters. This facilitates the alignment of the glucans within a cluster and of the latter with those of other clusters leading to hydrophobic structures that aggregate into insoluble material. These are packaged into a huge solid called the "starch granule." The major branched fraction of starch is called amylopectin which defines the largest biological molecule known. Amylose, the minor fraction of starch, can be considered as an accessory smaller and linear (with very few branches) side product within starch (Buléon et al. 1998). Indeed, some organisms fail to accumulate this polysaccharide while building normal starch granules. The presence of amylose is due to the only enzyme which seems to be active within the semicrystalline polysaccharide matrix defined by the starch granule (all other enzymes being only active in the soluble state) (for a review of amylose biosynthesis, see Ball et al. 1998). The granule-bound starch synthase thus synthesizes amylose processively within the granule, thereby generating a linear polymer which is protected from the action of the soluble branching enzymes by the presence of the polysaccharide matrix. Starch is only found in Archaeplastida, cryptophytes, alveolates, and a subgroup of particular unicellular diazotrophic cyanobacteria (for a review, see Ball et al. 2011). All evidence points to the evolution of starch from the related glycogen metabolic pathways. The transition from glycogen to starch has happened several times independently: at least once in cyanobacteria, once in the common ancestor of Archaeplastida, and at least once in an alveolate ancestor.

At endosymbiosis two pathways of storage polysaccharide metabolism merged to generate the archaeplastidal starch metabolism network (Coppin et al. 2005;

Patron and Keeling 2005; Deschamps et al. 2008a). These consisted of the pathways of eukaryotic glycogen metabolism and that of cyanophycean starch metabolism. Prokaryotic and eukaryotic pathways of glycogen–starch metabolism both depend on polymerization of glucose through the so-called glycogen (starch) synthases. These glycosyl transferases transfer an activated glucose from either a purine (ADP-glucose) or pyrimidine (UDP-glucose) glucosyl-nucleotide to the nonreducing end of a growing α-1,4-linked chain (for a review of glycogen metabolism in eukaryotes, see Roach 2002; Wilson et al. 2010; for bacteria, see Preiss 1984; Wilson et al. 2010). These purine or pyrimidine nucleotide sugars are synthesized from either ATP or UTP and glucose-1-P through the corresponding nucleotide sugar pyrophosphorylases. Branches are introduced in both bacteria and eukaryotes through hydrolysis of a preexisting α-1,4 linkage within an α-1,4-linked chain and transfer of a segment of chain in α-1,6 position on a neighboring chain. In eukaryotes and bacteria glycogen (starch) phosphorylases are considered to be the major enzymes of glycogen breakdown. These enzymes use orthophosphate to break the outer α-1,4 linkage at the nonreducing end of the glucan chains, thereby releasing glucose-1-P. This energy-efficient mobilization recovers one of the two high-energy phosphate bonds that were used to synthesize the nucleotide sugar substrate from glucose-1-P and either UTP or ATP. Glycogen phosphorylase stops four residues away from the first branch it encounters and cannot break or bypass the α-1,6 linkage. Hence the complete mobilization of the glycogen particle requires further action of the so-called glycogen debranching enzymes. Bacteria use direct debranching enzymes which release the outer glycogen chains in the form of malto-oligosaccharides (MOS) which need to be further degraded by enzymes of MOS metabolism which include maltodextrin phosphorylases and α-1,4 glucanotransferases (disproportionating enzymes or bacterial amylomaltases) (for a review of bacterial MOS metabolism, see Boos and Shuman 1998). Eukaryotes use indirect debranching enzymes which display a complex mode of debranching yielding solely glucose. In addition to phosphorolysis, hydrolysis plays an important role in eukaryotic glycogen breakdown and possibly also in bacterial catabolism. Of particular relevance, eukaryotes with the noticeable exception of fungi and animals often contain an exohydrolase named β-amylase that produces β-maltose from the outer chains of glycogen and which like phosphorylase stops a few residues before the branch (Deschamps et al. 2008a; Ball et al. 2011). The β-maltose is then degraded through an α-1,4 glucanotransferase named dpe2 (or amylomaltase) which transfers glucose from maltose to the outer α-1,4 chains of glycogen or heteroglycans (for a review of dpe2 action in starch and glycogen metabolism, see Fettke et al. 2009). In addition to this eukaryotes may contain glucosidases or other hydrolases in selective compartments such as the lysosome or fungal vacuole where they access glycogen particles that have been redirected through autophagy-like mechanisms (Wang et al. 2001; see Wilson et al. 2010 and Roach 2002 for reviews).

When one compares the glycogen metabolism networks of bacteria and eukaryotes two major differences can thus be found. Most bacteria synthesize glycogen through the bacteria-specific metabolite ADP-glucose while eukaryotes

always use UDP-glucose. The second difference consists in the presence of direct debranching enzymes which, in bacteria only, tie glycogen to malto-oligosaccharide metabolism while eukaryotes seem devoid of cytosolic MOS-degrading enzymes other than the dpe2-type amylomaltases.

ADP-glucose is devoted to the synthesis of glycogen in bacteria while UDP-glucose is a substrate common to many distinct pathways in all living cells. Hence the synthesis of ADP-glucose through ADP-glucose pyrophosphorylase defines the first committed step of glycogen synthesis in bacteria and as such is finely tuned by allosteric regulations while it is glycogen synthase which is regulated in eukaryotes. Of particular relevance to our discussion is the cyanobacterial ADP-glucose pyrophosphorylase (AGPase) (for a review of ADP-glucose pyrophosphorylase structure and function, see Ballicora et al. 2003). In cyanobacteria the enzyme is tightly coupled to photosynthesis through not only its substrates (glucose-1-P and ATP) but also its allosteric activator (3-PGA) and inhibitor (orthophosphate). Hence cyanobacterial glycogen synthesis is finely tuned by the Calvin cycle status and the ATP to Pi ratio.

On the other hand, eukaryotic glycogen synthases and phosphorylases are in addition known to be the subject of activation or inhibition through kinase and phosphatase cascades. Indeed it was by studying the regulation of animal glycogen metabolism that protein kinases were discovered (Krebs 1983).

As we shall see, the above similarities and differences were used to establish the first biochemical connection between the cyanobiont and its host.

Starch Metabolism in Archaeplastida

Starch is usually thought of as a plastidial storage polysaccharide by plant biologists. In fact green plants and algae (Chloroplastida) are the only organisms accumulating storage polysaccharides in this organelle. Red algae and glaucophytes and all starch-storing secondary endosymbiosis derivatives accumulate starch in either the cytosol (alveolates) or the periplast (derived from the red alga cytosol) (reviewed in Ball et al. 2011). Hence plastidial starch metabolism can be considered as the exception rather than the rule. The enzyme network composition of Glaucophyta, Chloroplastida, and Rhodophyceae is summarized in Table 1, which also emphasizes the phylogenetic origin of these enzyme sequences. For comparison we have also listed the extant amoebal and cyanobacterial networks as examples reflecting the possible composition of the ancestral networks present in the future cyanobiont and eukaryotic host before endosymbiosis (see preceding paragraph for details). We have chosen a representative member of a particular group of cyanobacteria that we suspect to descend from the cyanobacterial clade that donated the plastid. This gene-rich diazotrophic yet unicellular group of cyanobacteria was the only clade not ruled out by the synapomorphies examined recently by Gupta (2009) within cyanobacteria. In addition it is the only group reported to contain the GBSSI gene (Deschamps et al. 2008a) which defines one of

Table 1 Storage polysaccharide enzyme networks

Activity	Cyanobacteria (*C. watsonii*)	Eukaryotes (*E. histolytica*)	Chloroplastida (*O. tauri*))	Rhodophyceae (*C. merolae*)	Glaucophyta (*C. paradoxa*)	Ancestor
ADP-glucose pyrophosphorylase	1	0	3	0	0	1
Starch synthase (ADPG) SSIII-SSIV	2	0	3	0	1	1
Starch synthase (ADPG) SSI-SSII	0	0	2	0	0	
Starch (glycogen) synthase (UDPG)	0	1	0	1	2-3 ?	1-2
GBSS I	1	0	1	(0)1	2	1
Branching enzyme	3	1	2	1	3	3
Isoamylase/GlgX	1	0	3	2	3	1
Indirect debranching enzyme	0	1	0	0	1	1
Pullulanase	0	0	1	1	0	N/A
Phosphorylase	2	2	3	1	2	2
Glucanotransferase (dpe1)	1	0	1	0	0	N/A
Transglucosidase (dpe2)	0	2	1	2	3	1
β-amylase	0	4	2	1	1	1
Glucan water dikinase (GWD)	0	0	3	1	2	1
Phosphoglucan water dikinase	0	0	2	-	1	1
sex4-laforin glucan phosphatase	0	1	1	1	1	1

The storage polysaccharide enzyme networks are displayed for *Entamoeba histolytica* (eukaryotic host glycogen metabolism reference), *Crocosphaera watsonii* (representative unicellular cyanobacterium of clade B chosen as the cyanobiont reference), *Ostreococcus tauri* (Chloroplastida), *Cyanidioschizon merolae* (Rhodophyceae), and *Cyanophora paradoxa* (Glaucophyta). Little variation in number and type of isoforms occurs within each group and other species of Chloroplastida or Rhodophyceae would have been as well suited as those chosen here. The numbers correspond to the different number of isoforms found for each type of activity. In the case of Rhodophyceae we have listed GBSSI as (0)1 despite its absence in the *C. merolae* genome since this important enzyme of amylose synthesis has been reported in Porphyridiales red algae (Shimonaga et al. 2007). Phylogenetic origin of the different enzymes is represented by the background color: light blue for cyanobacteria, pink for eukaryotes, violet for Chlamydiales, and yellow for diverse proteobacteria. Note that pullulanase is polyphyletic in red and green algae. The reconstruction displayed as "ancestor" makes an assumption of simple vertical inheritance from the common ancestor which holds for all monophyletic enzymes and minimizes the number of isoforms to the minimal number needed to explain the present distribution of genes in the three archaeplastidal lineages. We have represented SSI–SSII with a darker blue background as the latter do not exist in extant cyanobacteria but were probably generated post-endosymbiosis by duplication and diversification from a cyanobacterial GBSSI ancestral gene. Because

(continued)

Table 1 (continued)

GWD–PWD works in a background of a preexisting pathway of glycogen phosphorylation in eukaryotes, we consider these novel activities as host derived and represent them with a pink hatched background. The status represented in the ancestor column corresponds to the second stage depicted in Fig. 1b. At the onset of plastid endosymbiosis, it is thought that storage polysaccharides of the host cytosol were not defined by starch as suggested by Table 1 but by glycogen particles (Fig. 1a) as is presently the case for all eukaryotes groups unrelated to plastid endosymbiosis. This implies that at that early stage the domain fusions that generated the PWD–GWD group of enzymes and the mutations of the GlgX bacterial ancestor that transformed this enzyme into an isoamylase had not yet occurred. In turn, the GBSSI gene would have not yet been transferred, thereby postponing the later appearance of the SSI–SSII enzymes

the few cases where a cyanobacterial gene was used within the archaeplastidal starch metabolism network (see below). It also fulfills the requirement relatively to the presumed gene-rich and possibly diazotrophic nature of the ancestor (Deusch et al. 2008). All GBSSI containing cyanobacteria accumulate starch-like polymers and belong to subgroup V according to Honda et al. (1999) within clade B (Gupta 2009). GBSSI in both cyanobacteria and Archaeplastida is an enzyme solely active within the semicrystalline matrix of starch-like structures. In both cases it is responsible for the synthesis of amylose (see above). It displays little or no activity as a soluble enzyme. From these considerations we can deduce that the plastid donor was probably a starch accumulating cyanobacterium. Yet this property is entirely fortuitous as the transition to starch in the archaeplastidal cytosol seems to have occurred independently of this except for the presence of GBSSI.

The three distinct archaeplastidal biochemical starch metabolism networks illustrated in Table 1 essentially differ by two criteria: the first being the nature of the glycosyl nucleotide used for starch synthesis and the second consisting in the level of redundancy of the network (the number of enzyme forms for each step).

First, both Glaucophyta and Rhodophyceae accumulate starch from UDP-glucose in the cytosol (Nyvall et al. 2001; Plancke et al. 2008) while Chloroplastida synthesize plastidial starch from ADP-glucose (Lin et al. 1988; Zabawinski et al. 2001). Second, Rhodophyceae metabolize starch with a mere 11–12 genes while Glaucophyta use over 20 genes and Chloroplastida do so by using a minimum of 30 and often more than 40 genes. For the essential part this increase in complexity reflects an increase in redundancy of enzymes. For instance while only one branching enzyme, one soluble starch synthase, and one β-amylase are found in the rhodophycean network, three branching enzymes, four soluble starch synthases, and a minimum of three β-amylases are found in the green alga network. These redundancies in the green lineage stem from post-endosymbiosis gene duplications. The redundancies witnessed in the Glaucophyta are of a different nature. With the noticeable exception of direct debranching enzymes and dpe2 eukaryotic amylomaltases, most redundancies stem from pre-endosymbiosis duplications and reflect probably a greater diversity of eukaryotic glycogen metabolism enzymes present in the amoeba-like ancestor. In line with this observation, only Glaucophyta still contain the gene coding indirect debranching enzyme that must have been present in this ancestor before it engulfed the cyanobacterium.

Notwithstanding these differences the vast majority of common steps that are listed in Table 1 are controlled by enzymes which display a clear common phylogenetic origin (Coppin et al. 2005; Patron and Keeling 2005; Deschamps et al. 2008a) in agreement with Archaeplastida monophyly (Rodríguez-Ezpeleta et al. 2005). In addition, the Archaeplastida storage metabolism network has witnessed only one invention (a novel enzyme activity not found elsewhere): the glucan and phosphoglucan water dikinases (GWD and PWD) which are required to mobilize the crystalline structures of starch which would otherwise remain inaccessible to hydrosoluble enzymes (reviewed in Fettke et al. 2009). These enzymes phosphorylate the insoluble amylopectin crystals from the β-phosphate of ATP through a dikinase reaction. This renders the otherwise hydrophobic crystals more

accessible to attack by hydrosoluble enzymes of starch catabolism such as β-amylases. The "invention" of the GWD/PWD enzymes results from a fusion of a CBM20 (carbohydrate binding module 20) with a dikinase domain. It probably occurred in the host cytosol since the introduction of the phosphate also requires the hydrolysis of the latter through glucan phosphatases. CBM20-containing glucan phosphatases have been reported as enzymes preventing the accumulation of abnormal hyper-phosphorylated glycogen in animals (Tagliabracci et al. 2008). In particular, "laforin" was documented as such a phosphatase which when defective led to the pathological accumulation of lafora bodies (abnormal glycogen) in the brains of humans afflicted by "lafora disease." A related defective "SEX4" glucan phosphatase of Arabidopsis was reported to prevent normal starch accumulation in this model plant (Kotting et al. 2009). Interestingly this defect could be complemented by the introduction of the animal laforin gene (Gentry et al. 2007). Hence, invention of the GWD/PWD enzymes was greatly facilitated by the preexisting metabolism of phosphorylated glycogen in eukaryotes. However, no bacteria have ever been reported to accumulate phosphorylated glycogen and this pathway is presumed to be inexistent in these organisms and in particular in cyanobacteria. Hence, the routes of starch catabolism in cyanobacteria are entirely different relying possibly on distinct enzymes and mechanisms. If the cyanobiont was initially a starch accumulator it remains highly unlikely that the loss of a catabolic enzyme from the cyanobiont genome could be complemented by the targeting of a corresponding host protein (i.e., of eukaryotic phylogeny) to the evolving symbiont.

The archaeplastidal starch metabolism network differs from that of eukaryotic glycogen metabolism (exemplified by *Entamoeba histolytica* in Table 1) by two critical steps. The first is defined by the aforementioned GWDs and PWDs, and the second concerns the presence of direct debranching enzymes (named isoamylase) which are never found in storage polysaccharide metabolism of heterotrophic eukaryotes. In green plants and algae the direct debranching enzymes are responsible for generating the asymmetrical and ordered distribution of branches that generates the cluster structure of amylopectin which in turn is required for aggregation into solid semicrystalline structures. These two features are sufficient to explain the difference between glycogen and starch metabolism in eukaryotes.

Looking at the starch synthases present in the three archaeplastida lineages it appears that the major soluble starch synthases responsible for amylopectin synthesis display a phylogenetic origin in agreement with its substrate specificity. Hence the UDP-glucose-specific enzymes are clearly related to enzymes of glycogen synthesis of CAZy family GT5 that use UDP-glucose in several eukaryotic lineages (with the noticeable exception of fungi and animals which use a GT3 enzyme) while SSIII–IV of Chloroplastida are related to bacterial GT5 enzymes that use ADP-glucose. However, the source of the chloroplastidal SSIII–IV enzyme is likely to be chlamydial rather than cyanobacterial (Moustafa et al. 2008; Ball et al. 2013). The source for the SSI–SSII enzymes likely was a duplication of the GBSSI gene that occurred selectively in Chloroplastida while the source of GBSSI is distinctively cyanobacterial (Deschamps et al. 2008a; Ball et al. 2013).

If we further look at the phylogenetic origin of the enzymes (Table 1) it appears that glaucophytes contain a complete network of eukaryotic glycogen metabolism. Rhodophyceae only lack indirect debranching enzyme while Chloroplastida lack indirect debranching enzyme and the GT5 UDP-glucose-specific glycogen(starch) synthase. However and very importantly, the only contribution of cyanobacteria to the archaeplastidal network consists of GBSSI in all three lineages and of GBSSI and ADP-glucose pyrophosphorylases in Chloroplastida. Enzymes that were previously thought of as cyanobacterial, like SSIII, SSIV, DPE1 (D-enzyme), pullulanase, and isoamylase turned out to display very clear distinct bacterial origins. Most importantly, SSIII-IV and isoamylase originated most probably from chlamydial intracellular pathogens (Moustafa et al. 2008).

Reconstruction of the Ancient Network of Starch Metabolism

The current predominant view states that Archaeplastida are monophyletic and that plastid endosymbiosis defines a common ancestor for the whole group (Rodríguez-Ezpeleta et al. 2005). We can then assume that most genes (with the exception of pullulanase which is not monophyletic and dpe1 which was selectively transferred horizontally from an unknown bacterium to the Chloroplastida) will have been transmitted vertically from the common ancestor to the three Archaeplastida lineages. Therefore, one can reconstruct the minimal network of enzymes which had to be present in the common ancestor to explain the present distribution of genes encoding them in Glaucophyta, Rhodophyceae, and Chloroplastida. In order to do this we further minimized the number of isoforms to a single enzyme when we have good reasons to suspect that gene duplications and subfunctionalization have occurred post-endosymbiosis. The ancestral pathway reconstructed in Table 1 relies on both ADP-glucose and UDP-glucose. This ancestral pathway contains a complete set of cytosolic eukaryotic glycogen metabolism enzymes but lacks all but two cyanobacterial enzymes (ADP-glucose pyrophosphorylase and GBSSI). Because all three lineages display the same pattern of loss (with the exception of ADP-glucose pyrophosphorylase which was kept by Chloroplastida only) it is more parsimonious to suppose that the loss occurred once in the common ancestor at a very early stage. We believe this implies that the cyanobiont had lost the opportunity to metabolize storage polysaccharides and that the latter were only present in the ancestor's cytosol. Indeed, if the cyanobiont gene losses occurred at a very early stage the complementation of a missing enzyme by the supply of a corresponding host enzyme would have been problematic. First the major plastidial protein targeting machinery (TOC–TIC) may not have been yet routinely efficient and second, as mentioned earlier, prior to plastid endosymbiosis starch-storing cyanobacteria had evolved mechanisms of polysaccharide mobilization entirely different from those found in Archaeplastida or heterotrophic glycogen

accumulating eukaryotes. This would make complementation of gene loss by supply of corresponding eukaryotic enzymes unlikely. The suggestion that the ancestor had lost the ability to store glucose in the cyanobiont is further strengthened by three additional observations.

First, Henrissat et al. (2002) have noted that parasites and symbionts in general tended to lose storage polysaccharide metabolism as a function of their obligatory intracellular lifestyle. This seems to be the case for the only other photosynthetic cyanobacterial endosymbiont reported: that carried by *Paulinella chromatophora* where the chromatophore genome seems to have already lost the genes of storage polysaccharide metabolism (Nowack et al. 2008).

Second, if storage polysaccharide metabolism was lost at a very early stage this would imply that the ancestor of Chloroplastida synthesized starch in the cytosol and the pathway was redirected to plastids when the green algae evolved. We have previously reviewed in detail the possible reasons and the problems dealing with the redirection of starch metabolism to plastids (Deschamps et al. 2008b, c). Suffice it to say here that this probably defined two intermediate stages (MOS and glycogen accumulation) that likely generated a requirement for duplications of enzyme forms followed by enzyme subfunctionalizations (Deschamps et al. 2008c). Indeed, the Chloroplastida have selectively experienced such duplication and subfunctionalization rounds leading to their characteristic highly redundant pathway (see above). This further suggests that the ancestral network was, as proposed initially, exclusively cytosolic.

Third, Glaucophyta and Rhodophyceae still synthesize starch exclusively in the cytosol today with no evidence for the presence of plastidial storage polysaccharides. Since Glaucophyta in general are assumed to have conserved a greater number of ancestral features it would further support that cytosolic starch deposition was ancestral in Archaeplastida and that plastidial starch synthesis is derived.

Taken together, the loss of the vast majority of enzymes of cyanobacterial starch metabolism and the three aforementioned observations make a compelling case for an ancient localization of storage polysaccharides exclusively in the host cytosol shortly after endosymbiosis. This would imply that the ADP-glucose-specific starch synthase would have been active in the cytosol. This in turn would require the presence of ADP-glucose in this compartment. However, ADP-glucose is not synthesized by eukaryotes. Hence, one would be tempted to place the cyanobacterial ADP-glucose pyrophosphorylase in the host cytosol shortly after endosymbiosis, thereby generating the ancestral cytosolic dual substrate pathway. The problem with such a hypothesis is its lack of physiological relevance. Dual substrate pathways are not common in biochemistry and the advantage of producing ADP-glucose in addition to UDP-glucose in the host cytosol is anything but obvious. As noted previously, ADP-glucose pyrophosphorylase in cyanobacteria and plants is an enzyme finely tuned by photosynthesis and the Calvin cycle. Because these processes have never moved out of the plastid, it seems reasonable to assume that ADP-glucose pyrophosphorylase will have never left this compartment during the whole evolution process. If the enzyme is left within the future

plastid stroma then the system to become functional requires the presence of an ADP-glucose transporter on the cyanobiont inner membrane to feed the cytosolic ADP-glucose requiring glycogen (starch) synthase. With such a transporter the physiological relevance of the reconstructed dual substrate pathway becomes enlightening. We have outlined above that storage can be predicted to define an interesting buffer between the unsynchronized supply and demand for carbon during plastid endosymbiosis. The proposed reconstruction of storage polysaccharide metabolism in the common ancestor of Archaeplastida lends considerable support to this prediction. Upon close examination of the carbon flux generated in this proposed reconstruction, it seems impossible to imagine a better suited first connection between the unrelated partners of plastid endosymbiosis. The carbon that flows through ADP-glucose pyrophosphorylase within the cyanobiont would have normally been committed to storage since this nucleotide sugar is devoted to glycogen (starch) synthesis in bacteria. Hence this carbon committed to leave cyanobacterial metabolism by becoming temporarily unavailable in the form of solid cyanophycean starch will similarly escape the latter by physically moving out of the cyanobiont through the ADP-glucose translocator. The cyanobacterial pathways have been optimized to generate and control this escape for millions of years and there are no penalties to be expected for such an export of carbon. This to our knowledge would not be the case for any other possible carbon substrate. Upon arrival in the host cytosol ADP-glucose is unlikely to affect host metabolism which does not recognize it. ADP-glucose will thus exclusively feed cytosolic glycogen synthesis and increase the available storage carbohydrate pools. The only very modest penalty will be the uncontrolled increase in the osmotic impact of glycogen which can be considered as negligible. Access to the cytosolic glycogen pools will be through the eukaryotic glycogen catabolism regulatory networks that have been tailored by millions of years of selection to respond optimally to the various needs of this ancient protist in a changing environment. The system was thus optimal at the very moment phagocytosis aborted and the connection was established. Hence, reconstruction of starch metabolism in the archaeplastida ancestor opens an unexpected window on the nature of the biochemical connection that drove plastid endosymbiosis. It should be emphasized here that reconstruction of starch metabolism in the common ancestor does not exactly reflect the situation present at the time endosymbiosis was established but rather shortly thereafter but before the three Archaeplastida lineages diverged. Indeed, in our reconstruction proposal starch is present in the cytosol and the GWD–PWD pathway of starch catabolism also. This situation is very close to that found in extant glaucophytes. It is however unlikely that the transition from glycogen to starch occurred immediately as this required the simultaneous "invention" of the dikinase-CBM20 gene fusion to generate the GWD–PWD required to catabolize the glucan crystals and the recruitment of a modified bacterial debranching enzyme to generate the "crystals." Indeed the bacterial source of this enzyme did not display the required "isoamylase" type of substrate specificity to begin with. However the latter is thought to be needed to synthesize crystalline amylopectin. The bacterial ancestors displayed a much narrower substrate specificity consisting of hydrolysis of external glycogen chains

of three to four glucose residues long (Dauvillée et al. 2005). The required modifications will have required gene duplications and the accumulation of mutations changing the enzyme specificity on the duplicated locus. Similarly the enzymes that work downstream from the phosphorylated crystals had to accumulate mutations in their genes that optimized their action by comparison to their previous analogous role in glycogen breakdown. This all took time and suggests that the transition came about later. It nevertheless happened fast enough to allow for the recruitment of the cyanobacterial GBSS gene by EGT which was otherwise likely to have been very quickly lost since the cyanobiont presumably became starchless very early on. The early glycogen and late starch accumulation stages are displayed in Fig. 1a, b. The reconstructed pathway from Table 1 reflects the late starch accumulation stage (Fig. 1b).

Incompatibility of Phototrophy and Diazotrophy in Photosynthetic Eukaryotes

As stated above, cyanobacterial metabolism has been tailored by natural selection to allow for escape of excess carbon in the form of ADP-glucose. Yet, we must still admit that it has also been selected to be able to tap carbon in the polysaccharide stores when required and this will certainly happen mostly in darkness. Hence the tolerance for carbon escape does not necessarily mean that cyanobacteria can do with no carbon stores. Mutants of storage polysaccharide synthesis completely lacking glycogen and starch, respectively, have been obtained in yeast (Thon et al. 1992), *E. coli* (Damotte et al. 1968), *Chlamydomonas reinhardtii* (Zabawinski et al. 2001), *Arabidopsis thaliana* (Lin et al. 1988), and more recently in both *Synechocystis* sp. PCC 6803 (Miao et al. 2003) and *Synechococcus elongatus* (Suzuki et al. 2010). In yeast, *E. coli*, and *Chlamydomonas* there is virtually no impact on growth of these microorganisms under laboratory conditions. In *Arabidopsis thaliana*, starchless mutants grow normally under continuous light. However, under day and night cycles growth of the mutant plants becomes stunted. In cyanobacteria, mutants lacking glycogen have been produced that carry a defect for the single cyanobacterial ADP-glucose pyrophosphorylase subunit (Miao et al. 2003; Suzuki et al. 2010). Growth of these mutants was monitored under continuous illumination. Under these conditions growth proved to be normal and the requirement for glycogen pools minimal. However, photosynthesis and respiration were impacted, with a significant reduction in photosynthesis especially under high light and a 50 % reduction in respiration activity in darkness. The reduction in photosynthesis was thought to be due to a decrease in the regeneration of oxidized $NADP^+$ allowed through glycogen synthesis. The decreased respiration was attributed to the fact that glycogen breakdown accounted for a significant portion of the accessible carbohydrate substrate pools.

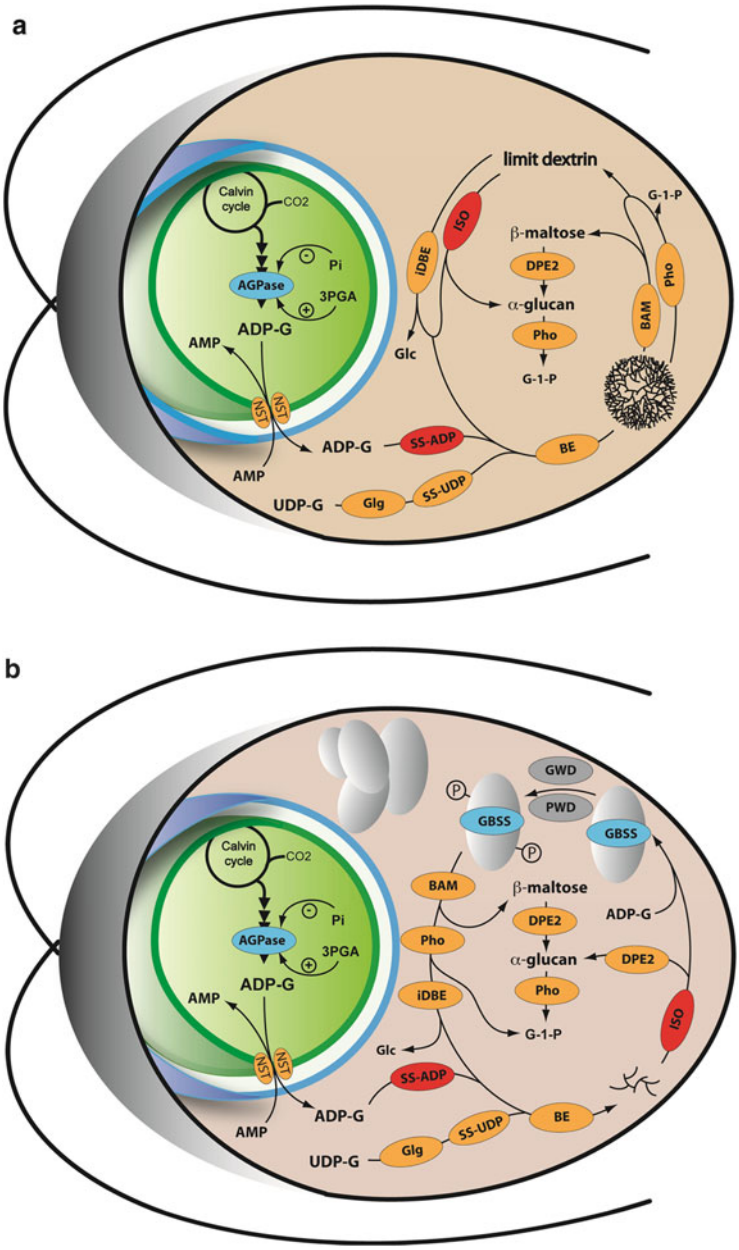


Fig. 1 (**a**) The storage polysaccharide network at the onset of endosymbiosis. In this reconstruction the transition to starch in the common ancestor cytosol has not yet occurred. This implies that the direct debranching enzyme (labeled *iso*) of bacterial phylogeny did not yet duplicate to generate the isoamylase required for amylopectin crystallization. As a consequence there was no need yet for the evolution of the GWD–PWD dikinases. The bacterial direct debranching enzyme still displays its ancestral bacterial function which is to debranch the product of glycogen

Quite interestingly, the *Synechoccus* mutants were shown to be also more sensitive to salt and photooxidative stresses. Indeed it was hypothesized by Deschamps et al. (2008c) that ATP in darkness may define a critical limitation in the absence of stored carbon. Under these conditions magnesium chelatase would not be able to assemble, thereby leading to the accumulation of photoactive intermediates of chlorophyll synthesis which upon the return of light would induce photooxidative stresses (Reinhold et al. 2007). It was in addition hypothesized by Deschamps et al. (2008c) that the selective increase in the chlororophyll synthesis that could have accompanied the evolution of chlorophyll *b*-containing antennae may have defined the selection pressure that has propelled the return of storage polysaccharides to the chloroplasts of the evolving green lineage.

Fig. 1 (continued) degradation (labeled *limit dextrin*) by glycogen phosphorylase (labeled *pho*). The eukaryotic indirect debranching enzyme has a similar function. However unlike the eukaryotic indirect debranching enzyme (idBE), the bacterial enzyme releases the maltotetraose outer chains (labeled *α-glucan*) in the cytosol which may have been subjected to degradation by a combination of the dpe2 amylomaltase and the glycogen phosphorylase. The cytosolic dual substrate pathway of glycogen accumulation relies on both UDP-glucose (UDP-G) generated through host biochemical networks in its cytosol according to host needs and ADP-glucose generated by the cyanobacterial ADP-glucose pyrophosphorylase (AGPase) which is activated by 3-PGA and inhibited by orthophosphate according to the cyanobiont's networks and physiology. To be incorporated into cytosolic glycogen this substrate has to be exported by a nucleotide sugar translocator (labeled *NST*) of host origin which exchanges ADP-glucose with AMP. The ADP-glucose substrate in the cytosol will have to be incorporated through an ADP-glucose-specific glucan synthase (labeled *SS-ADP*). On the other hand, the host UDP-glucose pools will be directed to glycogen according to the highly regulated eukaryotic UDP-glucose-specific glucan synthase (labeled *SS-UDP*). This enzyme unlike the bacterial glucan synthase requires a primer to elongate a glucan. This primer is defined by glycogenin, an autoglucosylating protein (labeled *Glg*). The glucans elongated through both glucan synthases will then be branched into glycogen by branching enzyme (labeled *BE*). The glycogen outer chains will be degraded through either β-amylase (labeled *BAM*) or glycogen phosphorylase (labeled *Pho*) to generate maltose and glucose-1-P, respectively. The maltose will be metabolized by the dpe2 amylomaltase. Enzymes of host phylogenetic origin are colored in *beige*. Those of cyanobacterial origin in *blue* and those of chlamydial origin in *red*. (**b**) Cytosolic storage polysaccharide metabolism has been reconstructed as detailed in the text. This early stage corresponds to the common ancestor starch metabolism after the transition from glycogen to starch has occurred. This transition required the duplication and evolution of the bacterial direct debranching enzyme into a functional isoamylase (iso). This enzyme processes the branches generated randomly on the hydrophilic branched polysaccharides generated by branching enzymes. The debranched glucans (labeled *α-glucan*) will be metabolized through a combination of amylomaltases and phosphorylases. Simultaneously a gene fusion between a CBM20 (carbohydrate binding module) possibly from laforin (see text) and a dikinase domain enabled the phosphorylation and loosening of the otherwise undegradable amylopectin crystals (displayed by the *circled P* attached to the white starch granules). This fusion generated the archaeplastidal GWD–PWD inventions (labeled in *gray*) which were required to initiate starch catabolism through the β-amylase and phosphorylases (see above). The presence of polysaccharides aggregated into semicrystalline starch granules enabled the binding and function of the cyanobacterial GBSS (displayed bound to starch) responsible for amylose synthesis within the polysaccharide matrix

In face of these results and speculations, we can predict how the loss of storage polysaccharide would have impacted the cyanobiont at endosymbiosis. We believe that unlike the cyanobacterial mutants, there would have been no impact at all on photosynthesis since the cyanobiont is still able to consume reducing equivalents through the synthesis and export of ADP-glucose in a fashion entirely similar to glycogen or starch synthesis of wild-type algae. In addition the cyanobiont being sheltered within an eukaryotic cytosol, we do not believe that it would have still required a particular resistance to osmotic stress. Yet we expect that respiration in darkness would have been dramatically reduced. Hence, there may have been a strong requirement for an ability to import ATP in darkness which was exacerbated by the need to assemble magnesium chelatase to obviate photooxydative stresses (Reinhold et al. 2007; Deschamps et al. 2008c). This may define the reason why all three Archaeplastida lineages have recruited ATP/ADP transport proteins (NTT, nucleotide transporter) on the inner membrane of their plastids which drive the unidirectional import of ATP in exchange for ADP. The gene encoding these transporters was acquired by lateral gene transfer from a Chlamydiale source (Linka et al. 2003). This transporter may have defined a critical early requirement for successful endosymbiosis. Although import of cytosolic ATP may have been sufficient to obviate photooxydative stresses we believe this import was unlikely to have allowed the maintenance of diazotrophy in the cyanobiont. Indeed, the cyanobacterial ancestor may be affiliated to extant unicellular diazotrophic cyanobacteria of clade B. A number of studies have suggested that these cyanobacteria reach the anoxia status required for nitrogen fixation in darkness thanks to the respiration of their large starch pools (Colón-López et al. 1997; Schneegurt et al. 1994). Not only was the vast amount of ATP and reducing power needed to feed nitrogenase but above all the consumption of the local O_2 by respiration would by itself be required. When the cyanobiont lost its storage polysaccharide pools very early on (possibly even at the onset of endosymbiosis) it lost the ability to fix nitrogen at the same time and thus very quickly also lost the *nif* genes. Indeed, while the NTT transporter could in theory supply the ATP to obviate photooxidative stresses it could not compensate for the decrease in the respiration activity absolutely required to reach anoxia. We believe this explains why in nature no member of the Archaeplastida has retained the ability to fix nitrogen that was originally displayed by the cyanobiont's ancestor. This also suggests that after endosymbiosis the cyanobiont had to be provided with some form of reduced nitrogen by the host.

The ADP-Glucose Connection

As mentioned above, the incorporation into glycogen of glucose from ADP-glucose in the common ancestor's cytosol required the transport of this glycosyl-nucleotide out of the cyanobiont. It is unlikely that the cyanobiont encoded such a transporter. Indeed the physiological significance of a protein exporting ADP-glucose into the

extracellular medium would be hard to imagine for free-living bacteria. Likewise the host is not likely to harbor such a transporter since ADP-glucose is neither synthesized nor used by eukaryotes. Clues to the elusive origin of the putative ancient ADP-glucose translocator came when Weber et al. (2006) (see also Facchinelli and Weber 2013) examined the phylogeny of the major extant plastidial carbon translocators. These belong to a family of proteins known as the pPT (phosphosugar phosphate translocator) proteins which exchange sugar phosphates for orthophosphate. The best studied transporter of this family is the TPT or triose phosphate translocator which is responsible for the export of carbon from the chloroplast to the plant leaf cell cytosol. Weber et al. (2006) demonstrated that the whole family of transporters found in green and red algae and secondary endosymbiosis derivatives was monophyletic. It was further assumed to have originated through duplication and evolution of a gene for a host endomembrane transporter. The authors postulated that this ancient transporter probably was involved in establishing the symbiotic flux.

Upon looking closer to the eukaryotic origin of these transporters these appeared to consist of members of a family of nucleotide sugar translocators (for a review, see Handford et al. 2006) known as the NST3 family (defined in Martinez-Duncker et al. 2003). Among the nucleotide sugar translocator families, NST3 defines a family that transports not only pyrimidine sugar nucleotides but also many purine sugar nucleotides (Martinez-Duncker et al. 2003). Interestingly, NST3 contains many GDP-mannose translocators, the latter defining a structural analog of ADP-Glc. Colleoni et al. (2010) thus tested the abilities of GDP-mannose translocators from yeast and *Arabidopsis* to transport ADP-Glc in yeast membrane-derived liposomes. They were able to show that the *Arabidopsis* GDP-mannose translocator was able to transport ADP-Glc as efficiently as GDP-mannose but displayed a lower affinity for the non-physiological substrate. Nevertheless, the Km for ADP-glucose remained at a 1–5 mM concentration range which is in agreement with a possible role of such a transporter in establishing the initial symbiotic link. Indeed, it is expected that mutants which are blocked in the utilization of ADP-glucose will see the size of their ADP-glucose pools rise above 1 mM as was demonstrated in cereal endosperm mutants (Shannon et al. 1996). The cyanobiont, having lost the ability to polymerize glucans but not to synthesize the nucleotide sugar substrate, was in precisely that situation.

One of the obvious problems faced by the host-endomembrane-derived putative ancestral ADP-glucose translocator was how to reach the cyanobiont's inner membrane at a time when no plastidial protein targeting machinery was likely to have existed. An interesting observation was published by Loddenkötter et al. (1993) concerning the expression in yeast of the TPT deprived of its transit peptide sequence. The protein was found to be located on the yeast mitochondrial membranes. Although contamination of ER membranes could not be definitively ruled out, the authors reported that in vitro also the protein was associated with yeast or Neurospora mitochondria in an energy- and receptor-independent fashion, strongly suggesting that this protein displayed an innate ability to reach the organelle membranes in the absence of a functional transport system. If such a property

was displayed by the ancestral transporter it would have greatly facilitated its recruitment at the onset of plastid endosymbiosis.

We thus postulate that such a transporter accidentally reached the cyanobiont's inner membrane. A duplicated gene encoding this transporter subsequently enabled it to be expressed and regulated independently of the endomembrane sugar nucleotide translocators. Later, evolution further explored the numerous possibilities of substrate exchange offered by this family of transporters whose expression had been optimized with respect to photosynthate export. This yielded, thanks to other duplicated copies, the pPT family that allowed for a more integrated solution to the export of carbon from plastids. Both the ADP-glc translocator and the pPT coexisted until the Archaeplastida lineages diverged. This happened when the emerging Archaeplastida lost the ability to synthesize glucans from ADP-glc in the cytosol. In Rhodophyceae, it happened when the ADP-glucose-specific starch synthase was lost while in Chloroplastida it happened at the final stages of rewiring of the storage polysaccharide metabolism network to plastids, i.e., when starch disappeared from the cytosol.

ADP-Glucose Transport and Glucan Polymerization

The second condition that had to be met at the onset of plastid endosymbiosis was the presence in the host cytosol of a glucan synthase able to use ADP-glc. The eukaryotic glycogen synthases use UDP-glc with little or no activity for purine nucleotide sugars as substrates. Immediate establishment of the symbiotic flux of carbon was however required to allow for natural selection of plastid endosymbiosis. This did not give the required time for the accumulation and selection of mutations in the host glucan transferase gene. Clearly an efficient ADP-glucose utilizing glycogen/starch synthase had to be present in the host cytosol at the onset of the event. Such enzymes are never observed in eukaryotes and have only been described in the bacterial or archean domains (with the exception of course of green algae and plants). To get further insights into this problem we have examined the phylogeny of extant archaeplastidal starch synthases that use ADP-glc (Ball et al. 2013). Two monophyletic groups are found in green algae and plants (Chloroplastida): the GBSSI–SSI–SSII group and the SSIII–IV group (Deschamps et al. 2008a; Ball et al. 2013). The GBSSI–SSI–SSII clade can be reasonably rooted by considering that the enzyme source is defined by the cyanobacterial GBSSI gene. With such a root in mind, the published phylogenetic trees support a transfer of the GBSSI gene to the Archaeplastida ancestor before the three lineages diverged. When GBSSI was bound to starch in the ancestor's cytosol it was exposed to the presence of both UDP-glc and ADP-glc that drove the ancient dual substrate pathways of storage polysaccharide synthesis. The gene therefore accumulated mutations that turned this low-affinity cyanobacterial enzyme that originally only used ADP-glc into a bifunctional synthase accepting both glycosyl nucleotides as

substrates to achieve amylose synthesis whenever either the ADP-glc or the UDP-glc pools rose above the required levels.

Under this scenario with cyanobacteria at the root of the clade, the GBSS1 gene duplicated and accumulated mutations that turned the duplicated gene product into a soluble (unbound) activity. This probably happened selectively in the Chloroplastida lineage as the pathway was redirected to plastids. Hence SSI–SSII were not available at the time of endosymbiosis and no bacterial glucan synthase shows significant proximity to these enzymes despite the presence of hundreds of available bacterial whole genome sequences (Ball et al. 2013). Under this hypothesis, which is consistent with the phylogeny and a cyanobacterial source of GBSSI, we can conclude that only wild-type GBSSI could have been available in the host cytosol. However this can only be imagined, provided an LGT had just happened shortly before or at endosymbiosis because of the phagotrophic habit of the ancestral protist ("you are what you eat"). This otherwise nonproductive event could thus have been selected to establish the symbiotic flux. However, this hypothesis does not stand in face of the biochemical properties of GBSSI. GBSSI is an enzyme that displays very little activity when expressed as soluble protein both in vivo (Dauvillée et al. 1999) and in vitro as a recombinant enzyme unbound to starch (Edwards et al. 1999). Yet at the time of endosymbiosis, the eukaryote ancestor synthesized glycogen and not starch which evolved shortly thereafter. A wild-type GBSSI protein would not have been able in such a context to polymerize glucan from ADP-glucose onto glycogen.

SSIII–SSIV presently define the only extant archaeplastidal transferases whose ancestor could have played a role in supplying the symbiotic link. One of the most surprising findings of the recently established *Cyanophora paradoxa* genome sequence (Price et al. 2012) was the description of an SSIII–SSIV-like enzyme sequence which is presumably involved in cytosolic starch synthesis. We believe this enzyme still uses ADP-glc only in present-day *Cyanophora*, but this yet needs to be demonstrated. This could thus suggest that *Cyanophora paradoxa* might very well define a living fossil of the putative ancestral dual substrate pathway, although it is apparently lacking both ADP-glucose pyrophosphorylase and the ADP-glucose translocator. In line with this suggestion, the GBSSI of glaucophytes displays similar affinities for both nucleotide sugars and the *C. paradoxa* phosphorylase is surprisingly exquisitely sensitive to mixed inhibition by ADP-glucose (Plancke et al. 2008). The source of the ADP-glucose in glaucophytes needs to be ascertained, but it is already known that the reversible sucrose synthase reaction using ADP in place of UDP is not involved since these organisms lack sucrose metabolism altogether (Price et al. 2012). What would be the rationale for glaucophytes to have kept this enzyme in its cytosol while they have lost the ability to produce the ADP-glucose substrate in plastids? Clues can be found in the exceptional properties displayed by the SSIII–SSIV starch (glycogen) synthases. These enzymes are involved in the priming of starch granules and thus control their numbers and sizes (Roldán et al. 2007; Szydlowski et al. 2009). In addition, unlike other starch synthase mutant combinations a double SSIII–SSIV defective mutant abolishes starch synthesis. This essential in vivo function correlates with the in vitro

ability displayed at least by SSIII to prime polysaccharide synthesis (Szydlowski et al. 2009). The GT3 UDP-glc requiring glycogen synthase of fungi and animals requires the presence of glycogenin, an autoglucosylating protein used as a primer (Cheng et al. 1995). The *Cyanophora paradoxa* genome contains no convincing glycogenin candidate sequence. Hence, the GT5 UDP-glucose requiring glycogen (starch) synthase of glaucophytes may have become dependent on the SSIII–SSIV-like enzyme for polysaccharide synthesis priming.

The SSIII–SSIV clade defines a monophyletic group consisting of the plant enzymes as well as a number of related enzymes from cyanobacteria, proteobacteria, and Chlamydiales (Ball et al. 2013). The phylogeny of this group is complex because of the existence of several LGT events splitting the Chlamydiales into two groups. Despite this complexity, all possible scenarios reject the cyanobacteria as the source of the archaeplastidal enzyme (Ball et al. 2013). The presence of the Chlamydiales at the base of the clade and the fact that the pathogens define the only group of organisms containing this glucan synthase as sole enzyme used for glycogen metabolism suggest that the enzyme may have evolved and acquired its exceptional biochemical properties in the pathogens. These genes were then passed on to Archaeplastida, proteobacteria, and cyanobacteria.

Clearly, the ancestor of extant SSIII–SSIV does qualify as a serious candidate to provide the enzyme used to establish the symbiotic link between the cyanobiont and its host. Why would such an enzyme have been present in the host cytosol at the advent of plastid endosymbiosis? We believe this question may be answered when a clear understanding of glycogen metabolism function will be reached in the group of organisms that are the most likely source for the LGT to Archaeplastida: the Chlamydiales intracellular pathogens (Moustafa et al. 2008).

References

Ball S, van de Wal M, Visser R (1998) Progress in understanding the biosynthesis of amylose. Trends Plant Sci 3:462–467

Ball SG, Colleoni C, Cenci U, Raj JN, Tirtiaux C (2011) The evolution of the glycogen and starch pathway in eukaryotes gives molecular clues to understand the establishment of plastid endosymbiosis. J Exp Bot 62:1775–1801

Ball SG, Subtil A, Bhattacharya D, Moustafa A, Weber APM, Gehre L, Colleoni C, Arias MC, Cenci U, Dauvillée D (2013) Metabolic effectors secreted by bacterial pathogens; essential facilitators of plastid endosymbiosis? Plant Cell 25(1):7–21

Ballicora MA, Iglesias AA, Preiss J (2003) ADP-glucose pyrophosphorylase, a regulatory enzyme for bacterial glycogensynthesis. Microbiol Mol Biol Rev 67:213–225

Boos W, Shuman H (1998) Maltose/maltodextrin system of *Escherichia coli*: transport, metabolism, and regulation. Microbiol Mol Biol Rev 62:204–229

Buléon A, Colonna P, Planchot V, Ball S (1998) Starch granules: structure and biosynthesis. Int J Biol Macromol 23:85–112

Chan XC, Gross J, Yoon HS, Bhattacharya D (2011) Plastid origin and evolution: new models provide insights into old problems. Plant Physiol 155:1552–1560

Cheng C, Mu J, Farkas I, Huang D, Goebl MG, Roach PJ (1995) Requirement of the self-glucosylating initiator proteins Glg1p and Glg2p for glycogen accumulation in *Saccharomyces cerevisiae*. Mol Cell Biol 15:6632–6640

Colleoni C, Linka M, Deschamps P, Handford MG, Dupree P, Weber APM, Ball SG (2010) Phylogenetic and biochemical evidence supports the recruitment of an ADP-glucose translocator for the export of photosynthate during plastid endosymbiosis. Mol Biol Evol 27:2691–2701

Colón-López MS, Sherman DM, Sherman LA (1997) Transcriptional and translational regulation of nitrogenase in light-dark- and continuous-light-grown cultures of the unicellular cyanobacterium Cyanothece sp. strain ATCC 51142. J Bacteriol 179:4319–4327

Coppin A, Varre JS, Lienard L, Dauvillée D, Guerardel Y, Soyer-Gobillard MO, Buléon A, Ball S, Tomavo S (2005) Evolution of plant-like crystalline storage polysaccharide in the protozoan parasite *Toxoplasma gondii* argues for a red alga ancestry. J Mol Evol 60:257–267

Damotte M, Cattanéo J, Sigal N, Puig J (1968) Mutants of Escherichia coli K 12 altered in their ability to store glycogen. Biochem Biophys Res Commun 32:916–920

Dauvillée D, Colleoni C, Shaw E, Mouille G, D'Hulst C, Morell M, Samuel MS, Bouchet B, Gallant DJ, Sinskey A, Ball S (1999) Novel starch-like polysaccharides are synthesized by a soluble form of granule-bound starch synthase in glycogen accumulating mutants of *Chlamydomonas reinhardtii*. Plant Physiol 119:321–330

Dauvillée D, Kinderf IS, Li Z, Kosar-Hashemi B, Samuel MS, Rampling L, Ball S, Matthew MK (2005) Role of the *E. coli glgX* gene in glycogen metabolism. J Bacteriol 187:1465–1473

Deschamps P, Colleoni C, Nakamura Y, Suzuki E, Putaux JL, Buléon A, Haebel S, Ritte G, Steup M, Falcon LI, Moreira D, Loeffelhardt W, Nirmal Raj J, Plancke C, D'Hulst C, Dauvillée D, Ball S (2008a) Metabolic symbiosis and the birth of the plant kingdom. Mol Biol Evol 25:536–548

Deschamps P, Moreau H, Worden AZ, Dauvillée D, Ball SG (2008b) Early gene duplication within chloroplastida and its correspondence with relocation of starch metabolism to chloroplasts. Genetics 178:2373–2387

Deschamps P, Haferkamp I, D'Hulst C, Neuhaus E, Ball S (2008c) The relocation of starch metabolism to chloroplasts: when, why and how. Trends Plant Sci 13:1802–1816

Deusch O, Landan G, Roettger M, Gruenheit N, Kowallik KV, Allen JF, Martin W, Dagan T (2008) Genes of cyanobacterial origin in plant nuclear genomes point to a heterocyst-forming plastid ancestor. Mol Biol Evol 25:748–761

Edwards A, Borthakur A, Bornemann S, Venail J, Denyer K, Waite D, Fulton D, Smith A, Martin C (1999) Specificity of starch synthase isoforms from potato. Eur J Biochem 266:724–736

Facchinelli F, Weber APM (2013) Insertion of metabolite transporters into the endosymbiont membrane(s) as a prerequisite for primary endosymbiosis. In: Löffelhardt W (ed) Endosymbiosis. Springer, Wien New York, pp 53–80

Fettke J, Hejazi M, Smirnova J, Höchel E, Stage M, Steup M (2009) Eukaryotic starch degradation: integration of plastidial and cytosolic pathways. J Exp Bot 60:2907–2922

Gentry MS, Dowen RH, Worby CA, Mattoo S, Ecker JR, Dixon JE (2007) The phosphatase laforin crosses evolutionary boundaries and links carbohydrate metabolism to neuronal disease. J Cell Biol 178:477–488

Gupta RS (2009) Protein signatures (molecular synapomorphies) that are distinctive characteristics of the major cyanobacterial clades. Int J Syst Evol Microbiol 59:2510–2526

Handford M, Rodriguez-Furlán C, Orellana A (2006) Nucleotide-sugar transporters: structure, function and roles *in vivo*. Braz J Med Biol Res 39:1149–1158

Henrissat B, Deleury E, Coutinho PM (2002) Glycogen metabolism loss: a common marker of parasitic behaviour in bacteria? Trends Genet 18:437–440

Honda D, Yokota A, Sugiyama J (1999) Detection of seven major evolutionary lineages in cyanobacteria based on the 16S rRNA gene sequence analysis with new sequences of five marine *Synechococcus* strains. J Mol Evol 48:723–739

Kotting O, Santelia D, Edner C, Eicke S, Marthaler T, Gentry MS, Comparot-Moss S, Chen J, Smith AM, Steup M, Ritte G, Zeeman SC (2009) STARCH-EXCESS4 is a laforin-like phosphoglucan phosphatase required for starch degradation in *Arabidopsis thaliana*. Plant Cell 21:334–346

Krebs EG (1983) Historical perspectives on protein phosphorylation and a classification system for protein kinases. Philos Trans R Soc B Biol Sci 302:3–11

Lin T-P, Caspar T, Somerville C, Preiss J (1988) Isolation and characterisation of a starchless mutant of *Arabidopsis thaliana* (L.) Heynh. lacking ADP-glucose pyrophosphorylase activity. Plant Physiol 86:1131–1135

Linka N, Hurka H, Lang BF, Burger G, Winkler HH, Stamme C, Urbany C, Seil I, Kusch J, Neuhaus HE (2003) Phylogenetic relationships of non-mitochondrial nucleotide transport proteins in bacteria and eucaryotes. Gene 306:27–35

Loddenkötter B, Kammerer B, Fischer K, Flügge UI (1993) Expression of the functional mature chloroplast triose phosphate translocator in yeast internal membranes and purification of the histidine-tagged protein by a single metal-affinity chromatography step. Proc Natl Acad Sci USA 90:2155–2159

Martinez-Duncker I, Mollicone R, Codogno P, Oriol R (2003) The nucleotide-sugar transporter family: a phylogenetic approach. Biochimie 85:245–260

Miao X, Wu Q, Wu G, Zhao N (2003) Changes in photosynthesis and pigmentation in an *agp* deletion mutant of the cyanobacterium *Synechocystis* sp. Biotechnol Lett 25:391–396

Moustafa A, Reyes-Prieto A, Bhattacharya D (2008) Chlamydiae has contributed at least 55 genes to Plantae with predominantly plastid functions. PLoS One 3:e2205

Nakamura Y, Takahashi J, Sakurai A, Inaba Y, Suzuki E, Nihei S, Fujiwara S, Tsuzuki M, Miyashita H, Ikemoto H, Kawachi M, Sekiguchi H, Kurano N (2005) Some cyanobacteria synthesize semi-amylopectin type alpha-polyglucans instead of glycogen. Plant Cell Physiol 46:539–545

Nowack ECM, Melkonian M, Glöckner G (2008) Chromatophore genome sequence of *Paulinella* sheds light on acquisition of photosynthesis by eukaryotes. Curr Biol 18:410–418

Nyvall P, Pelloux J, Davies HV, Pedersen M, Viola R (2001) Purification and characterization of a novel starch synthase selective for uridine 5′-diphosphate glucose from the red alga *Gracilaria tenuistipitata*. Planta 209:143–152

Patron NJ, Keeling PK (2005) Common evolutionary origin of starch biosynthetic enzymes in green and red algae. J Phycol 41:1131–1141

Plancke C, Colleoni C, Deschamps P, Dauvillée D, Nakamura Y, Haebel S, Ritte G, Steup M, Buléon A, Putaux JL, Dupeyre D, D'Hulst C, Ral JP, Löffelhardt W, Ball SG (2008) The pathway of starch synthesis in the model glaucophyte *Cyanophora paradoxa*. Eukaryot Cell 7:247–257

Preiss J (1984) Bacterial glycogen synthesis and its regulation. Annu Rev Microbiol 38:419–458

Price DC, Chan CX, Yoon HS, Yang EC, Qiu H, Weber APM, Schwacke R, Gross J, Blouin NA, Lane C, Reyes-Prieto A, Durnford DG, Neilson JAD, Lang BF, Burger G, Steiner JM, Löffelhardt W, Meuser JE, Posewitz MC, Ball S, Arias MC, Henrissat B, Coutinho PM, Rensing SA, Symeonidi A, Doddapaneni H, Green BR, Rajah VD, Boore J, Bhattacharya D (2012) *Cyanophora paradoxa* genome elucidates origin of photosynthesis in algae and plants. Science 335:843–847

Raven JA, Beardall J, Flynn KJ, Maberly SC (2009) Phagotrophy in the origins of photosynthesis in eukaryotes and as a complementary mode of nutrition in phototrophs: relation to Darwin's insectivorous plants. J Exp Bot 60:3975–3987

Reinhold T, Alawady A, Grimm B, Beran KC, Jahns P, Conrath U, Bauer J, Reiser J, Melzer M, Jeblick W, Neuhaus HE (2007) Limitation of nocturnal import of ATP into Arabidopsis chloroplasts leads to photooxidative damage. Plant J 50:293–304

Roach PJ (2002) Glycogen and its metabolism. Curr Mol Med 2:101–120

Rodríguez-Ezpeleta N, Brinkmann H, Burey SC, Roure B, Burger G, Löffelhardt W, Bohnert HJ, Philippe H, Lang BF (2005) Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes. Curr Biol 15:1325–1330

Roldán I, Wattebled F, Lucas MM, Delvallé D, Planchot V, Jiménez S, Pérez R, Ball S, D'Hulst C, Mérida A (2007) The phenotype of soluble starch synthase IV defective mutants of *Arabidopsis thaliana* suggests a novel function of elongation enzymes in the control of starch granule formation. Plant J 49:492–504

Schneegurt MA, Sherman DM, Nayar S, Sherman LA (1994) Oscillating behavior of carbohydrate granule formation and dinitrogen fixation in the cyanobacterium Cyanothece sp. strain ATCC 51142. J Bacteriol 176:1586–1597

Shannon JC, Pien FM, Liu KC (1996) Nucleotides and nucleotide sugars in developing maize endosperms: synthesis of ADP-glucose in brittle-1. Plant Physiol 110:835–843

Shearer J, Graham TE (2002) New perspectives on the storage and organization of muscle glycogen. Can J Appl Physiol 27:179–203

Shimonaga T, Fujiwara S, Kaneko M, Izumo A, Nihei S, Francisco BP, Satoh A, Fujita N, Nakamura Y, Tsuzuki M (2007) Variation in storage alpha-polyglucans of red algae: amylose and semi-amylopectin types in Porphyridium and glycogen type in Cyanidium. Mar Biotechnol 9:192–202

Suzuki E, Ohkawa H, Moriya K, Matsubara T, Nagaike Y, Fujiwara S, Tsuzuki M, Nakamura Y (2010) Carbohydrate metabolism in mutants of the cyanobacterium *Synechococcus elongatus* PCC 7942 defective in glycogen synthesis. Appl Environ Microbiol 76:3153–3159

Szydlowski N, Ragel P, Raynaud S, Lucas MM, Roldán I, Montero M, Muñoz FJ, Ovecka M, Bahaji A, Planchot V, Pozueta-Romero J, D'Hulst C, Merida A (2009) Starch granule initiation in *Arabidopsis* requires the presence of either class IV or class III starch synthases. Plant Cell 21:2443–2457

Tagliabracci VS, Girard JM, Segvich D, Meyer C, Turnbull J, Zhao X, Minassian BA, Depaoli-Roach AA, Roach PJ (2008) Abnormal metabolism of glycogen phosphate as a cause for Lafora disease. J Biol Chem 283:33816–33825

Thon VJ, Vigneron-Lesens C, Marianne-Pepin T, Montreuil J, Decq A, Rachez C, Ball SG, Cannon JF (1992) Coordinate regulation of glycogen metabolism in the yeast *Saccharomyces cerevisiae*: induction of glycogen branching enzyme. J Biol Chem 267:15224–15228

Wang Z, Wilson WA, Fujino MA, Roach PJ (2001) Antagonistic controls of autophagy and glycogen accumulation by Snf1p, the yeast homolog of AMP-activated protein kinase, and the cyclin-dependent kinase Pho85p. Mol Cell Biol 21:5742–5752

Weber APM, Linka M, Bhattacharya D (2006) Single, ancient origin of a plastid metabolite translocator family in Plantae from an endomembrane-derived ancestor. Eukaryot Cell 5:609–612

Wilson WA, Roach PJ, Montero M, Baroja-Fernández E, Muñoz FJ, Eydallin G, Viale AM, Pozueta-Romero J (2010) Regulation of glycogen metabolism in yeast and bacteria. FEMS Microbiol Rev 34:952–985

Zabawinski C, Van Den Koornhuyse N, D'Hulst C, Schlichting R, Giersch C, Delrue B, Lacroix JM, Preiss J, Ball S (2001) Starchless mutants of *Chlamydomonas reinhardtii* lack the small subunit of a heterotetrameric ADP-glucose pyrophosphorylase. J Bacteriol 183:1069–1077

Analysis of the Genome of *Cyanophora paradoxa*: An Algal Model for Understanding Primary Endosymbiosis

Debashish Bhattacharya, Dana C. Price, Cheong Xin Chan, Jeferson Gross, Jürgen M. Steiner, and Wolfgang Löffelhardt

Abstract Algae and plants rely on the plastid (e.g., chloroplast) to carry out photosynthesis. This organelle traces its origin to a cyanobacterium that was captured over a billion years ago by a single-celled protist. Three major photosynthetic lineages (the green algae and plants [Viridiplantae], red algae [Rhodophyta], and Glaucophyta) arose from this primary endosymbiotic event and are putatively united as the Plantae (also known as Archaeplastida). Glaucophytes comprise a handful of poorly studied species that retain ancestral features of the cyanobacterial endosymbiont such as a peptidoglycan cell wall. Testing the Plantae hypothesis and elucidating glaucophyte evolution has in the past been thwarted by the absence of complete genome data from these taxa. Furthermore, multigene phylogenetics has fueled controversy about the frequency of primary plastid acquisitions during

D. Bhattacharya (✉)
Rutgers University, Department of Ecology, Evolution and Natural Resources, Institute of Marine and Coastal Science, 59 Dudley Road, Rutgers University, New Brunswick, NJ 08901, USA
e-mail: bhattacharya@aesop.rutgers.edu

D.C. Price • J. Gross
Department of Ecology, Evolution and Natural Resources, Rutgers University, New Brunswick, NJ 08901, USA

C.X. Chan
Department of Ecology, Evolution and Natural Resources, Rutgers University, New Brunswick, NJ 08901, USA

The University of Queensland, Institute for Molecular Bioscience and ARC Centre of Excellence in Bioinformatics, Brisbane, QLD 4072, Australia

J.M. Steiner
Department of Plant Physiology, Martin-Luther-University Halle-Wittenberg, 06120 Halle (Saale), Germany

W. Löffelhardt
Max F. Perutz Laboratories, Department of Biochemistry and Cell Biology, University of Vienna, 1030 Vienna, Austria

W. Löffelhardt (ed.), *Endosymbiosis*, DOI 10.1007/978-3-7091-1303-5_7,

© Springer-Verlag Wien 2014

eukaryote evolution because these approaches have generally failed to recover Plantae monophyly and often provide conflicting results. Here, we review some of the key insights about Plantae evolution that were gleaned from a recent analysis of a draft genome assembly from *Cyanophora paradoxa* (Glaucophyta). We present results that conclusively demonstrate Plantae monophyly. We also describe new insights that were gained into peptidoglycan biosynthesis in glaucophytes and the carbon concentrating mechanism (CCM) in *C. paradoxa* plastids.

Contents

Introduction 136
Genome Data 137
Testing Plantae Monophyly 138
Enzymes of Peptidoglycan Biosynthesis 141
The Rubisco-Containing Microcompartment of Muroplasts: Carboxysome Versus Pyrenoid 144
Conclusions 146
References 146

Introduction

Independent primary endosymbioses involving Gram-negative bacteria led to the origin of the bioenergetic organelles mitochondria and plastids (e.g., Margulis 1970; Cavalier-Smith and Lee 1985; Gray 1992; Gross and Bhattacharya 2009). These events had lasting impacts on our planet with plastid endosymbiosis giving rise to algae and plants that became a driving force behind Earth's climate, geochemistry, and ecology (Falkowski et al. 2004). Primary plastids are shared by three extant lineages that are referred to as the Plantae (Cavalier-Smith 1981) or Archaeplastida (Adl et al. 2005): the Glaucophyta (glaucophyte algae), the Rhodophyta (red algae), and the Viridiplantae (green algae and plants). Determining the number of primary endosymbioses (i.e., single or multiple) that gave rise to the plastid in these three phyla has long been an open question in algal evolution and directly impacts inference of the tree of life. If it happened as many as three times in the different Plantae lineages that would suggest that these taxa are polyphyletic and establishing the combination of a eukaryotic and a prokaryotic cell is relatively "easy," in evolutionary terms. Most importantly, the resulting chimera converged on similar plastid features for each separate occurrence. If it occurred only once, then the plastid in all major autotrophic eukaryotic lineages traces its origin to this remarkable event in evolution.

A large body of data stemming from phylogenetic and comparative analyses of plastids suggest that primary endosymbiosis occurred a single time in the Plantae ancestor and all plastids [except in the *Paulinella* lineage (e.g., Yoon et al. 2006; Nowack et al. 2011)] trace their ancestry to this singular event (e.g., Bhattacharya et al. 2004; Delwiche 1999; Palmer 2003; Rodriguez-Ezpeleta et al. 2005;

Chan et al. 2011). Nevertheless, many recent nuclear multigene trees provide little (Burki et al. 2007; Patron et al. 2007) or no (Nozaki et al. 2009; Baurain et al. 2010; Parfrey et al. 2010) support for Plantae monophyly and often provide conflicting results. The inability to conclusively support or reject Plantae monophyly (and thereby resolve the number of plastid primary endosymbioses in the eukaryote tree of life) is largely explained by the lack of complete genome data from glaucophytes. This hurdle was recently crossed with the completion and analysis of a draft genome assembly from the glaucophyte *Cyanophora paradoxa* (Price et al. 2012). Intriguingly, the *C. paradoxa* plastid (often referred to as the "muroplast") maintains the ancestral cyanobacterial trait of a peptidoglycan wall (Pfanzagl et al. 1996). This and other traits such as an unconventional carbon-concentrating mechanism (CCM) have made this glaucophyte a model for photosynthesis and endosymbiosis research. Here, we review some of the major features of the *C. paradoxa* genome and the insights it provides into Plantae evolution.

Genome Data

To generate an initial genome draft, a total of 8.3 billion base pairs (Gbp) of Roche 454 and Illumina GAIIx sequence data from *C. paradoxa* CCMP329 (Pringsheim strain) were coassembled with 279 Mbp of random-shear Sanger sequence from this taxon. The resulting assembly comprised 60,119 contigs totaling 70.2 Mbp in size with an N50 of 2.7 kbp (minimum 100 bp and maximum 66 kbp). This highly fragmented assembly is currently being improved by the addition of significant Illumina mate-pair library sequence data. Pulsed field gel electrophoresis analysis of *C. paradoxa* shows the presence of at least seven chromosomes with the smallest being less than 3 Mbp in size (Price et al. 2012). A previous fluorescence-activated cell sorting (FACS) study suggested the haploid genome size in *C. paradoxa* was 140 Mbp (Löffelhardt et al. 1997). Given that the draft genome assembly converged on ca. 70 Mbp, it is likely that the original FACs sorting was done with diploid cells and the haploid genome size of *C. paradoxa* is closer to 70 Mbp.

Analysis of the *C. paradoxa* nuclear genome data (Price et al. 2012) demonstrates a highly enriched G + C-content (i.e., 83.8 % at third codon positions), resulting in difficulties in sequence generation and assembly. To assess whether the assembly was deficient in coding regions due to this issue, we used BLASTN with 3,900 Sanger-derived EST unigenes from *C. paradoxa* to query the draft assembly. This analysis showed that 99.1 % of the ESTs (i.e., putative protein coding regions) had hits (at $e \leq 10^{-10}$). This suggests the majority of expressed genes are represented in the genome data. Thereafter, a total of 15 Gbp of Illumina mRNA-seq data was used to train ab initio gene predictors, resulting in 27,921 weighted consensus gene structures. The organelle genomes of *C. paradoxa* were also analyzed, including novel data from its sister glaucophyte, *Glaucocystis nostochinearum*, but will not be discussed here (for details, see Price et al. 2012).

Testing Plantae Monophyly

Several approaches were taken to ascertain support for the monophyly of Plantae. These included phylogenetic analysis of single proteins, elucidating the extent and type of endosymbiotic or horizontal gene transfer (E/HGT), comparative analysis of groups of proteins such as components of the plastid protein translocons and fermentation pathways, and analysis of plastid solute transporters. All of these data strongly support a single origin of Plantae and therefore a single primary plastid endosymbiosis in their common ancestor (Price et al. 2012). Here, we present the results of the single protein and plastid translocon analyses.

For the single proteins, we first used BLASTP to analyze the evolutionary affiliations of the 27,921 predicted protein models in *C. paradoxa*. A total of 4,628 proteins had significant BLASTP hits ($e \leq 10^{-10}$) to sequences in a comprehensive local database that we use for comparative analysis (e.g., Moustafa et al. 2009; Chan et al. 2011). A simplified reciprocal BLAST best-hits approach (Chan et al. 2011) identified a total of 606 proteins that had hits only to one other phylum (i.e., exclusive gene sharing). With the requirement of an increasing number of hits per query (x) from *C. paradoxa* and the second phylum, as $x \geq 2$ (606 proteins), $x \geq 10$ (125 proteins), and $x \geq 20$ (23 proteins), we found that *C. paradoxa* and Viridiplantae shared the largest number of exclusive genes (Fig. 1), indicating a close evolutionary relationship between these lineages. Next, using an automated approach (Chan et al. 2011), we generated 4,445 maximum likelihood (ML) trees for *C. paradoxa* proteins that had significant database hits. To minimize the impact of taxon sampling on this analysis, we considered trees that contained $\geq$3 phyla and a minimum number of terminal taxa (N) that ranged from 4 to 40 (Fig. 2a). Using this approach, we found that >60 % of all trees support (at bootstrap $\geq$90 %) a sister group relationship between glaucophytes and red and/or green algae. The glaucophytes were most often positioned as sister to Viridiplantae (105, 83, 48, 19, and 10 trees at $N = 4$, 10, 20, 30, and 40), consistent with the analysis of exclusive gene sharing with only a small number of trees favoring the monophyly of glaucophytes and red algae. This result was found even though a significant number of trees favored glaucophyte–red–green (Plantae) monophyly (44, 40, 32, 18, and 16 trees at $N = 4$, 10, 20, 30, and 40) and we had substantial red algal genome data in our database (361,625 sequences). Interestingly, many of the trees showed *C. paradoxa* to be monophyletic with other Plantae in a clade ("shared") that also included non-Plantae phyla (GlR/GlGr/GlRGr in Fig. 2a). When we sorted the phylogenomic output using the red or green algae as the query to test Plantae monophyly, these results also identified Plantae as the most frequently recovered clade (Fig. 2b, c). However, both red and green algae show far more gene sharing than glaucophytes because they, unlike glaucophytes, are implicated in secondary endosymbioses that have resulted in their genes being spread throughout the tree of life to groups such as "chromalveolates" and euglenids (Harper and Keeling 2003; Moustafa et al. 2009; Baurain et al. 2010; Chan et al. 2011). Given that single protein

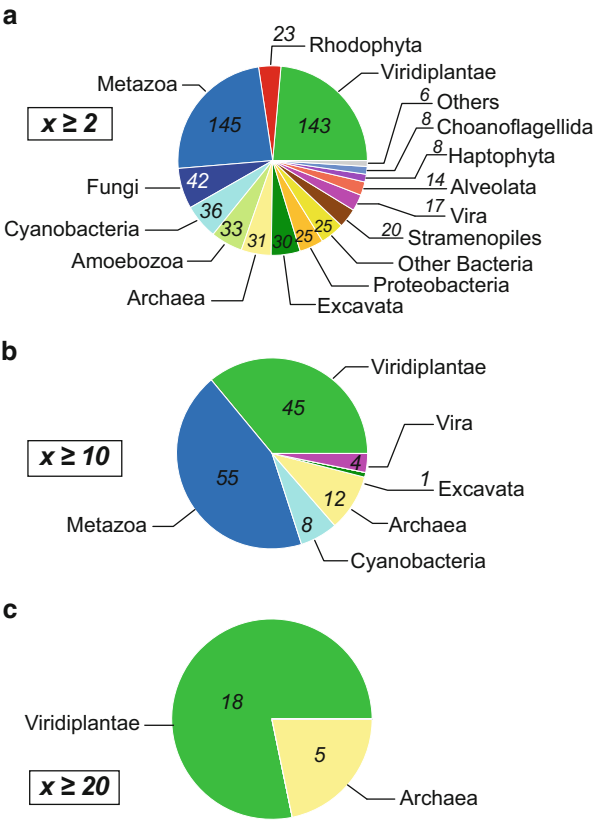


Fig. 1 Exclusive gene sharing between Glaucophyta and one other taxon, when the total number of hits (x) was $\geq$2 (**a**), $\geq$10 (**b**), and $\geq$20 (**c**). The *green slices* get larger as x increases. The vira matches include prasinophyte (*Bathycoccus*, *Micromonas*, and *Ostreococcus*) and *Chlorella* viruses

trees firmly establish glaucophytes as members of the Plantae we analyzed a landmark trait of Plantae, the plastid protein translocons.

A key innovation required for the cyanobacterium-to-plastid evolutionary transition in primary endosymbiosis was the establishment of protein translocons for protein targeting into the emergent organelle (e.g., Reumann et al. 2005; Gross and Bhattacharya 2008, 2009). Components of the *T*ranslocons at the *o*uter and *i*nner envelope membranes of *c*hloroplasts (Toc and Tic, respectively) have been described in higher plants, and algae of the green, red, and "chromalveolate" lineages (McFadden and van Dooren 2004). The existence of an analogous protein import system in *C. paradoxa* is suggested by immunological detection of epitopes in this alga using plant Toc75 and Tic110 antibodies, and heterologous protein import assays (Steiner et al. 2005; Yusa et al. 2008). These data suggest that all Plantae share a key invention that laid the foundation for plastid integration within the host cell. Our analysis of the *C. paradoxa* genome identified homologs of Toc75 and Tic110 that are OEM (outer envelope membrane) and IEM (inner envelope membrane) protein conducting channels, respectively, two Toc34-like receptors, as well as homologs of the plastid Hsp70 and Hsp93 chaperones, and stromal

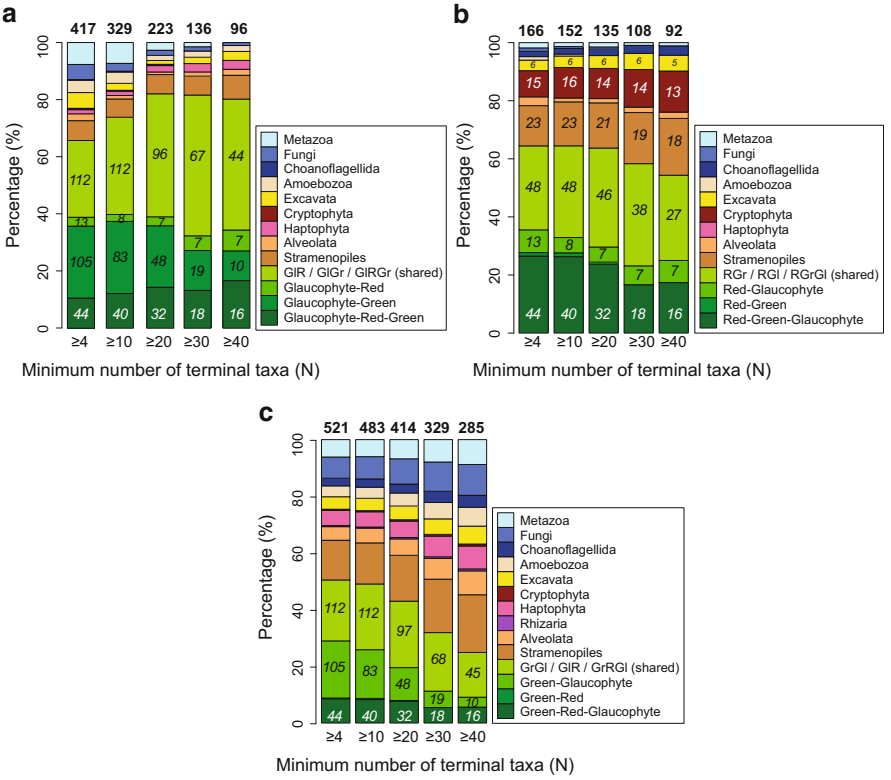


Fig. 2 Testing Plantae monophyly. (**a**) Percentage of maximum likelihood single protein trees that support the monophyly of Glaucophyta (bootstrap ≥90 %) with other members of the Plantae, or in combination with non-Plantae taxa that interrupt this clade. These latter groups of trees are primarily explained by red or green algal endosymbiotic gene transfer (EGT) into the nuclear genome of "chromalveolates" and euglenids. For each of these algal lineages, the set of trees with different numbers of terminal taxa (N) ≥4, ≥10, ≥20, ≥30, and ≥40 and distinct phyla ≥3 in a tree are shown. Similar analyses that were done using red algae and Viridiplantae as the query to test for Plantae monophyly are shown in panels (**b**) and (**c**), respectively

processing peptidase (Price et al. 2012). Such a minimal set of components is likely to have formed the primitive protein translocation system in the Plantae ancestor (Gross and Bhattacharya 2008, 2009). Candidates for additional translocon subunits were also detected in *C. paradoxa*. Furthermore, the Tic20 and Toc22 ML phylogenies provide unambiguous evidence for a cyanobacterial provenance of these genes in Plantae and a monophyletic relationship of *C. paradoxa* with plants and other algae (see Price et al. 2012). In summary, the evolution of protein translocons to the nascent plastid has long been held as a formative event in the emergence of the Plantae ancestor. Analysis of the *C. paradoxa* genome revealed the presence of the conserved core of translocon subunits derived from the cyanobacterial endosymbiont (i.e., Toc75, Tic20, Tic22) as well as novel genes

that apparently evolved de novo in the host (i.e., Toc34 and Tic110). These data provide further unambiguous evidence that the primary plastid was established in a single common ancestor of the Plantae.

Enzymes of Peptidoglycan Biosynthesis

The muroplast wall of glaucophyte algae is the sole documented example of peptidoglycan (PG) in Plantae and its origin from the plastid endosymbiont is noncontroversial. The PG consists of one giant molecule ("sacculus") and belongs to the A1gamma type, like the cell walls of *Escherichia coli* and cyanobacteria, but is thicker and more cross-linked than in the former and more reduced than in the latter. A unique feature of the PG is its modification with *N*-acetylputrescine (Löffelhardt and Bohnert 2001). The space between the inner and outer envelope membranes of muroplasts, the "periplasmic space", harbors the peptidoglycan layer and enzymes for its synthesis, modification, and degradation.

PG biosynthesis has been studied in great detail in *E. coli* and can be divided into cytoplasmic, membrane-bound, and periplasmic steps. This three-step compartmentalization process is also present in *C. paradoxa*: (1) biosynthesis of the disaccharide–pentapeptide precursor occurs in the muroplast stroma (activities of MurA and MurF have been shown), (2) its transfer to the lipid carrier at the inner envelope membrane, and (3) its insertion into growing PG chains in the periplasmic space (Löffelhardt and Bohnert 2001). The latter step is catalyzed by penicillin binding proteins (PBPs) that have transglycosylase and/or transpeptidase activities. Seven PBPs that range in size from 35 to 110 kDa were identified in the muroplast envelope by labeling with a radioactive derivative of ampicillin. In addition, enzymatic activities of DD- and LD-carboxypeptidases and DD-endopeptidase that hydrolyze defined bonds in PG have been demonstrated in muroplasts (Löffelhardt and Bohnert 2001).

Here, three different approaches were used for PBP gene identification (1) domain search; (2) BLAST search against the eight PBP genes of *Synechocystis* sp. PCC6803 (Marbouty et al. 2009) and the *Anabaena* sp. PCC7120 homologs; and (3) BLAST search against *Physcomitrella patens* PBP-like genes. In most cases, the results converged leading to at least 11 genes or gene fragments being identified in *C. paradoxa* (Table 1). In general, sequence similarity was higher to homologs in cyanobacteria than those in *P. patens*. No *C. paradoxa* homologs to the small PBPs 6 and 7 were identified. The PBP numbering scheme applied here is from *E. coli*. However, sequence similarity (especially among the large PBPs) is significant which is reflected in their redundant function.

In some cases of periplasmic proteins, bipartite presequences consisting of a transit peptide and a signal peptide can be envisaged. This suggests import to the muroplast stroma, followed by export to the periplasmic space. This special variant of "conservative sorting" would necessitate a dual location of Sec (already

Table 1 Nuclear genes involved in the biosynthesis of plastid peptidoglycan in *C. paradoxa*

Gene/protein	Function[a]	Contig
PBP3, PBP1, PBP2	PG transglycosylase/transpeptidase	11,577
PBP3, PBP1, PBP2	PG transglycosylase/transpeptidase	55,323
PBP3, PBP1	PG transglycosylase/transpeptidase	10,395
ftsI/PBP4	PG transglycosylase/transpeptidase	11,029
PBP?	PG transpeptidase	15,041
PBP 8	D-Ala-D-Ala-carboxypeptidase C	40,415
dacB/PBP 5	D-Ala-D-Ala-carboxypeptidase D-Ala-D-Ala-endopeptidase	7,465
dacB/PBP5	D-Ala-D-Ala-carboxypeptidase D-Ala-D-Ala-endopeptidase	6,898
vanX	D-Ala-D-Ala-dipeptidase	54,463
vanY/endolysin	D-Ala-D-Ala-carboxypeptidase	15,693
Lysozyme-like	Muramidase	54,844
mlt	Lytic transglycosylase	9,267
glmS	Glucosamin-6-P synthase	39,008
murA	UDP-*N*-acetylglucosamine-1-carboxyvinyl transferase	53,904
murA	UDP-*N*-acetylglucosamine-1-carboxyvinyl transferase	7,908
murB	UDP-*N*-acetylenolpyruvoyl-glucosamine reductase	8,011
murB	UDP-*N*-acetylenolpyruvoyl-glucosamine reductase	40,266
murC	UDP-*N*-acetylmuramate:L-Ala ligase	17,182
murI	Glutamate racemase	25,539
murD	D-Glu adding enzyme	39,147
murE	DAP-adding enzyme	9,035
Alr	Alanine racemase	53,160; 386
Ddl	D-Ala:D-Ala ligase	7,167
murF	UDP-*N*-acetylmuramoyl tripeptide/D-Ala-D-Ala ligase	52,912
mraY	Lipid I synthesis	37,977
murG	Lipid II synthesis	54,819

[a]The high MW (1–4) and the medium MW (5–8) PBPs are redundant in *Synechocystis* sp. PCC6803

documented) and Tat (seems possible as another parallel to cyanobacteria) translocases on thylakoid and inner envelope membranes of muroplasts. In a Gram-negative background, the low molecular weight (MW) peptidases VanX and VanY are not linked to vancomycin resistance but rather to D-alanine recycling and to an additional endolysin, respectively. Peptidoglycan biosynthesis requires cleavage of existing glycan chains to allow for insertion of new material. This is

performed by soluble and membrane-bound lytic transglycosylases: one gene of this kind could also be identified in *C. paradoxa*. A lysozyme family protein with significant similarity to protist lysozymes displays a signal peptide indicating a vacuolar (lysosomal) location that is likely involved in the autophagosomal digestion of damaged muroplasts. Genes for stromal proteins that are involved in the synthesis of the soluble precursor are denoted as *glm*. The N-terminal transit peptide identifies one protein in *C. paradoxa* (*glmS*) as a member of the muroplast-resident PG biosynthesis pathway, whereas a cytosolic counterpart would be expected to participate in protein glycosylation. The complete list of enzymes in the alga that are involved in UDP-*N*-acetylmuramate biosynthesis as well as the peptide side-chain adding enzymes, and the alanine (Alr) and glutamate (MurI) racemases are listed in Table 1. The membrane-bound or associated MraY and MurG proteins complete this compilation.

Genes for enzymes of PG biosynthesis were transferred twice into Plantae in the course of evolution—from the mitochondrial ancestor and from the cyanobacterial ancestor of plastids. These remain recognizable in sequence from *Arabidopsis thaliana* (few genes) to the moss *P. patens* (almost complete set), but their functions are likely to have changed. As long as chemical and structural proof is lacking (pleiotropic), effects of antibiotics or gene knock-outs of plastid division do not provide sufficient evidence to claim the presence and biosynthesis of PG in the plastids of bryophytes (Takano and Takechi 2010). Glaucophyte PG is unique in Plantae. In *Paulinella*, the situation is different: there is also PG in this eukaryote, but all genes necessary for its biosynthesis (Marin et al. 2007) are encoded on the endosymbiont (i.e., "chromatophore", photosynthetic organelle) genome, which exceeds the size of plastid genomes by a factor of 5–10. Unlike their counterparts in *C. paradoxa*, these genes retain their prokaryotic character, i.e., they were not transferred to the nuclear genome and thus no import of precursor proteins is required for biosynthesis of the sacculus in photosynthetic *Paulinella* species.

The correlation of more than one gene to a given function is not uncommon among cyanobacteria. A second gene with high sequence similarity to *murG* is more closely related to MGDG synthases, the likely function of "MurG" in plants. In an analogous fashion, *murD*-like genes might instead play a role in folate biosynthesis. Until the presence of PG in *P. patens* is unequivocally proven, one should expect modified functions for "mur-like" genes. The fact that the cyanobacterial counterparts are often, but not always the top hits suggests a mosaic structure of the gene complement for PG biosynthesis in *C. paradoxa*. HGT from bacteria (e.g., Firmicutes and Verrucomicrobia) is likely to be prominent when the transferred genes provide a required function, i.e., PG biosynthesis in the case of glaucophytes. In addition, gene replacement might have occurred in some cases.

The Rubisco-Containing Microcompartment of Muroplasts: Carboxysome Versus Pyrenoid

The conspicuous, electron-dense central body of *C. paradoxa* muroplasts described in most publications was named a carboxysome (Raven 2003; Fathinejad et al. 2008). This coinage did not take into account the fact that eukaryotes contain pyrenoids to fulfill the function of a carbon-concentrating mechanism (CCM) and emphasized the often-postulated transitional position of glaucophytes between plastids and cyanobacteria. However, all of our attempts to identify carboxysomal shell proteins in the *C. paradoxa* genome failed, either with domain searches (BMC = Pfam 00936 or Pfam 03319) or with a concatenated dataset of cyanobacterial CcmKLMNO sequences. Indeed, it might be problematic to harbor shell protein genes in the nucleus, because they have high affinities to each other and likely self-assemble as carboxysomal prestructures (Kinney et al. 2011), thereby interfering with protein import into muroplasts. Thus far, *Paulinella* constitutes the only example of "eukaryotic carboxysomes". Again, the necessary genes remain on the plastid genome, interestingly derived via HGT (Marin et al. 2007). In any case, the hypothesis of peptidoglycan retention in *C. paradoxa* (Raven 2003) to stabilize the plastid against the osmotic pressure of bicarbonate that is enriched more than 1,000-fold in the stroma through the action of the carboxysomal CCM could not be verified. In contrast, evidence was obtained (Table 2) for a number of proteins (LciB, C, and D) with functions in the pyrenoidal CCM of *Chlamydomonas reinhardtii* (Yamano et al. 2010). LciB and LciC were shown to form a hexameric complex (ca. 360 kDa) under active operation of the CCM: light and low concentration of CO_2. This complex localizes close to the pyrenoid but is relocalized from the pyrenoid to the stroma upon high CO_2 concentration or darkness. There seems to be no connection to pyrenoid development and/or starch sheath formation. A role is assumed in trapping of CO_2 that has escaped from the pyrenoid via interaction with the carbonic anhydrase Cah6 and, eventually, also in accumulating CO_2 reaching the stroma from the cytosol, i.e., in the active uptake of CO_2 in *C. reinhardtii* (Wang et al. 2011). Alternatively, physical blockage of CO_2 from escaping the pyrenoid by the complex has been postulated (Yamano et al. 2010). The complex is not required under high levels of CO_2. In this case, a function similar to the cyanobacterial shell proteins CcmK and CcmL (which, however, are present under all conditions) can be envisaged. Some putative cyanobacterial plastid ancestors contain LciB and LciC, given their filamentous nature (*Lyngbya*) or capability of producing a starch-like reserve carbohydrate (*Cyanothece*). These bacteria might use mechanisms of the type discussed above that are superimposed on their carboxysomal CCM. If carboxysomes were transferred to early plastids via endosymbiosis, the separation between carboxysomal and pyrenoidal CCM could have occurred within the phylum Glaucophyta, i.e., *C. paradoxa* and *G. nostochinearum* already progressed towards a pyrenoidal CCM, whereas *Gloeochaete wittrockiana* and *Cyanoptyche gloeocystis*, with their polyhedral microcompartments confined by an electron-dense shell-like

Table 2 Genes for proteins involved in the CCM of *Cyanophora paradoxa*

Gene	Function	Contig	Comments
LciA[a]	Bicarbonate transport	8,717	TP, complete
LciA	Bicarbonate transport	53,293	TP, complete
LciB[a]	CCM	37,097	TP, complete
LciB	CCM	53,135	TP, complete
LciB, *LciD*?	CCM	54,037	TP, 3′-truncated
LciB, *LciC*?	CCM	25,875	Fragment
rca[a]	Rubisco activase	26,296	TP, complete
?[a]	Carbonic anhydrase	52,891	Beta-CA superfamily, cytosolic
Cah4[a]	Carbonic anhydrase	54,421	Beta-CA superfamily, mitochondrial
Cah5[a]	Carbonic anhydrase	9,670	Beta-CA superfamily, mitochondrial
?	Carbonic anhydrase	38,132	Gamma-CA family, cytosolic?
?	Carbonic anhydrase	53,783	Gamma-CA family, cytosolic

[a]CO_2-responsive gene
TP muroplast transit peptide

layer, might have retained the carboxysomal CCM (Fathinejad et al. 2008). Under such a scenario, the *ccmKLMNO* genes would be expected to reside on the muroplast genomes of *G. wittrockiana* and *C. gloeocystis*. The PG wall, though no longer necessary, was retained for unknown reasons in the plastids of *C. paradoxa* and *G. nostochinearum*. Table 2 includes two genes encoding the putative bicarbonate transporter LciA and several genes with strong sequence similarity to genes for LciB, LciC, and LciD from *C. reinhardtii*. Because these are closely related, an exact assignment is difficult. However, whenever the N-termini are intact, unequivocal muroplast presequences were found for these enzymes.

A key enzyme of the CCM is carbonic anhydrase, either copackaged with Rubisco in cyanobacterial carboxysomes or located in the lumen of thylakoids traversing the pyrenoid of *C. reinhardtii*. The number of CAs can vary among algae, e.g., from 9 in *C. reinhardtii* to 13 in some diatoms (Tachibana et al. 2011). Five CAs from *C. paradoxa* are shown in Table 2. Two of these belong to the gamma-CA family with high sequence similarity to homologs in plants. The other three contain the conserved Zn-binding site (VCGHSHCGAMKG) of (cyano) bacterial beta-CAs. In the case of the putative mitochondrial CAs, high sequence similarity to *C. reinhardtii* Ca1 and Ca2 is observed. A bona fide muroplast CA (e.g., the stromal Cah6 or the lumenal Cah3 of *C. reinhardtii*) is missing from this compilation. If we assume a pyrenoidal CCM in *C. paradoxa*, the organism must utilize a mechanism different from that in *C. reinhardtii*. There is no evidence of a thylakoid-lumenal CA or a muroplast microcompartment traversed by thylakoid membranes. In the diatom *Phaeodactylum tricornutum*, the carbonic anhydrase CA-1 (CO_2 responsive) is copackaged with pyrenoidal Rubisco and does not reside in the lumen of the traversing thylakoid (Tachibana et al. 2011). Mass spectrometric analysis of central body proteins from *C. paradoxa* did not reveal a CA-like protein. The only outcome of these studies (in addition to Rubisco LSU and SSU) was

Rubisco activase that was also corroborated by Western blotting and assembly studies after in vitro import into isolated muroplasts (Fathinejad et al. 2008). *C. paradoxa* activase, while showing high sequence similarity to both cyanobacterial and plant homologs, lacks the C-terminal extension typical for filamentous cyanobacteria. This protein contains a domain that shares high sequence similarity with repetitive regions found in the largest carboxysome shell protein CcmM. An N-terminal extension present in plant homologs is present in the *C. paradoxa* protein. Taken together, the domain structure of Rubisco activase from *C. paradoxa* does not support the carboxysome concept. Several genes listed in Table 2 were shown to be CO_2 responsive in the closely related *C. paradoxa* SAG 45.84 (Kies strain) underlining their postulated role in the CCM (Burey et al. 2007).

Conclusions

Rather than being a relict lineage, the analyses presented here and in Price et al. (2012) paint a picture of the "living fossil" *C. paradoxa* as a gene- (and function)-rich species that provides many clues to early events in plastid endosymbiosis and Plantae evolution. These data unambiguously support Plantae monophyly, thereby answering a fundamental question about the eukaryote tree of life. The components of the peptidoglycan biosynthetic pathway in *C. paradoxa* were identified and indicated a cyanobacterial provenance of many key enzymes with likely instances of recruitment of additional genes via HGT from other prokaryote sources. Finally, evidence was found that strongly argues against the existence of a proposed eukaryotic carboxysome in *C. paradoxa*. The available data are more consistent with a pyrenoidal CCM in this species and in its sister *G. nostochinearum*. However, the mechanism of CCM function is likely to be different from that found in *C. reinhardtii*.

References

Adl SM, Simpson AG, Farmer MA, Andersen RA, Anderson OR, Barta JR, Bowser SS, Brugerolle G, Fensome RA, Fredericq S, James TY, Karpov S, Kugrens P, Krug J, Lane CE, Lewis LA, Lodge J, Lynn DH, Mann DG, McCourt RM, Mendoza L, Moestrup O, Mozley-Standridge SE, Nerad TA, Shearer CA, Smirnov AV, Spiegel FW, Taylor MF (2005) The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. J Eukaryot Microbiol 52:399–451

Baurain D, Brinkmann H, Petersen J, Rodríguez-Ezpeleta N, Stechmann A, Demoulin V, Roger AJ, Burger G, Lang BF, Philippe H (2010) Phylogenomic evidence for separate acquisition of plastids in cryptophytes, haptophytes, and stramenopiles. Mol Biol Evol 27:1698–1709

Bhattacharya D, Yoon HS, Hackett JD (2004) Photosynthetic eukaryotes unite: endosymbiosis connects the dots. Bioessays 26:50–60

Burey SC, Poroyko V, Ozturk ZN, Fathi-Nejad S, Schüller C, Ohnishi N, Fukuzawa H, Bohnert HJ, Löffelhardt W (2007) Acclimation to low [CO2] by an inorganic carbon concentrating mechanism in *Cyanophora paradoxa*. Plant Cell Environ 30:1422–1435

Burki F, Shalchian-Tabrizi K, Pawlowski J (2007) Phylogenomics reshuffles the eukaryotic supergroups. PLoS ONE 2:e790

Cavalier-Smith T (1981) Eukaryote kingdoms: seven or nine? BioSystems 14:461–481

Cavalier-Smith T, Lee JJ (1985) Protozoa as hosts for endosymbioses and the conversion of symbionts into organelles. J Protozool 32:376–379

Chan CX, Yang EC, Banerjee T, Yoon HS, Martone PT, Estevez JM, Bhattacharya D (2011) Red-and-green algal monophyly and extensive gene sharing found in a rich repertoire of red algal genes. Curr Biol 21:328–333

Delwiche CF (1999) Tracing the thread of plastid diversity through the tapestry of life. Am Nat 154:S164–S177

Falkowski PG, Katz ME, Knoll AH, Quigg A, Raven JA, Schofield O, Taylor FJ (2004) The evolution of modern eukaryotic phytoplankton. Science 305:354–360

Fathinejad S, Steiner JM, Reipert S, Marchetti M, Allmaier G, Burey SC, Ohnishi N, Fukuzawa H, Löffelhardt W, Bohnert HJ (2008) A carboxysomal CCM in the cyanelles of the "coelacanth" of the algal world, *Cyanophora paradoxa*? Physiol Plant 133:27–32

Gray MW (1992) The endosymbiont hypothesis revisited. Int Rev Cytol 141:233–357

Gross J, Bhattacharya D (2008) Revaluating the evolution of the Toc/Tic protein translocons. Trends Plant Sci 14:13–20

Gross J, Bhattacharya D (2009) Mitochondrial and plastid evolution in eukaryotes: an outsider's perspective. Nat Rev Genet 10:495–505

Harper JT, Keeling PJ (2003) Nucleus-encoded, plastid-targeted glyceraldehyde-3-phosphate dehydrogenase (GAPDH) indicates a single origin for chromalveolate plastids. Mol Biol Evol 20:1730–1735

Kinney JN, Axen SD, Kerfeld CA (2011) Comparative analysis of carboxysome shell proteins. Photosynth Res 109:21–32

Löffelhardt W, Bohnert HJ (2001) The cyanelle (muroplast) of *Cyanophora paradoxa*: a paradigm for endosymbiotic organelle evolution. In: Seckbach J (ed) Symbiosis. Kluwer, Dordrecht, pp 111–130

Löffelhardt W, Bohnert HJ, Bryant DA (1997) The complete sequence of the *Cyanophora paradoxa* cyanelle genome (Glaucocystophyceae). Plant Syst Evol 11(Suppl):149–162

Marbouty M, Mazouni K, Saguez C, Cassier-Chauvat C, Chauvat F (2009) Characterization of the *Synechocystis* strain PCC 6803 penicillin-binding proteins and cytokinetic proteins FtsQ and FtsW and their network of interactions with ZipN. J Bacteriol 191:5123–5133

Margulis L (1970) Origin of eukaryotic cells. Yale University Press, New Haven, CT, p 349

Marin B, Nowack ECM, Glöckner G, Melkonian M (2007) The ancestor of the *Paulinella* chromatophore obtained a carboxysomal operon by horizontal gene transfer from a *Nitrococcus*-like gamma-proteobacterium. BMC Evol Biol 7:85

McFadden GI, van Dooren GG (2004) Evolution: red algal genome affirms a common origin of all plastids. Curr Biol 14:R514–R516

Moustafa A, Beszteri B, Maier UG, Bowler C, Valentin K, Bhattacharya D (2009) Genomic footprints of a cryptic plastid endosymbiosis in diatoms. Science 324:1724–1726

Nowack EC et al (2011) Endosymbiotic gene transfer and transcriptional regulation of transferred genes in *Paulinella chromatophora*. Mol Biol Evol 28:407–422

Nozaki H, Maruyama S, Matsuzaki M, Nakada T, Kato S, Misawa K (2009) Phylogenetic positions of Glaucophyta, green plants (Archaeplastida) and Haptophyta (Chromalveolata) as deduced from slowly evolving nuclear genes. Mol Phylogenet Evol 53:872–880

Palmer JD (2003) The symbiotic birth of plastids: how many times and whodunit? J Phycol 39:4–11

Parfrey LW, Grant J, Tekle YI, Lasek-Nesselquist E, Morrison HG, Sogin ML, Patterson DJ, Katz LA (2010) Broadly sampled multigene analyses yield a well-resolved eukaryotic tree of life. Syst Biol 59:518–533

Patron NJ, Inagaki Y, Keeling PJ (2007) Multiple gene phylogenies support the monophyly of cryptomonad and haptophyte host lineages. Curr Biol 17:887–891

Pfanzagl B, Zenker A, Pittenauer E, Allmaier G, Martinez-Torrecuadrada J, Schmid ER, De Pedro MA, Löffelhardt W (1996) Primary structure of cyanelle peptidoglycan of *Cyanophora paradoxa*: a prokaryotic cell wall as part of an organelle envelope. J Bacteriol 178:332–339

Price DC, Chan CX, Yoon HS, Yang EC, Qiu H, Weber APM, Schwacke R, Gross J, Blouin NA, Lane C, Reyes-Prieto A, Durnford DG, Neilson JAD, Lang BF, Burger G, Steiner JM, Löffelhardt W, Meuser JE, Posewitz MC, Ball S, Arias MC, Henrissat B, Coutinho PM, Rensing SA, Symeonidi A, Doddapaneni H, Green BR, Rajah VD, Boore J, Bhattacharya D (2012) *Cyanophora paradoxa* genome elucidates origin of photosynthesis in algae and plants. Science 335:843–847

Raven JA (2003) Carboxysomes and peptidoglycan walls of cyanelles: possible physiological functions. Eur J Phycol 38:47–53

Reumann S, Inoue K, Keegstra K (2005) Evolution of the general protein import pathway of plastids (review). Mol Membr Biol 22:73–86

Rodriguez-Ezpeleta N, Brinkmann H, Burey SC, Roure B, Burger G, Löffelhardt W, Bohnert HJ, Philippe H, Lang BF (2005) Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes. Curr Biol 15:1325–1330

Steiner JM, Yusa F, Pompe JA, Löffelhardt W (2005) Homologous protein import machineries in chloroplasts and cyanelles. Plant J 44:646–652

Tachibana M, Allen AE, Kikutani S, Endo Y, Bowler C, Matsuda Y (2011) Localization of putative carbonic anhydrases in two marine diatoms, *Phaeodactylum tricornutum* and *Thalassiosira pseudonana*. Photosynth Res 109:205–221

Takano H, Takechi K (2010) Plastid peptidoglycan. Biochim Biophys Acta 1800:144–151

Wang Y, Duanmu D, Spalding MH (2011) Carbon dioxide concentrating mechanism in *Chlamydomonas reinhardtii*: inorganic carbon transport and CO2 recapture. Photosynth Res 109:115–122

Yamano T, Tsujikawa T, Hatano K, Ozawa S, Takahashi Y, Fukuzawa H (2010) Light and low-CO2 -dependent LCIB–LCIC complex localization in the chloroplast supports the carbon-concentrating mechanism in *Chlamydomonas reinhardtii*. Plant Cell Physiol 51:1453–1468

Yoon HS, Reyes-Prieto A, Melkonian M, Bhattacharya D (2006) Minimal plastid genome evolution in the *Paulinella* endosymbiont. Curr Biol 16:R670–R672

Yusa F, Steiner JM, Löffelhardt W (2008) Evolutionary conservation of dual Sec translocases in the cyanelles of *Cyanophora paradoxa*. BMC Evol Biol 8:304

Part III
Recent “Primary” Endosymbioses

Photosynthetic *Paulinella*: Recapitulation of Primary Plastid Establishment

Hwan Su Yoon, Eun Chan Yang, Huan Qiu, and Debashish Bhattacharya

Abstract The origin of photosynthesis in eukaryotes stems from a single primary endosymbiosis between a heterotrophic protist cell and a cyanobacterium that occurred more than 1.5 billion years ago. This proto-algal population gave rise to three lineages of the Plantae (Rhodophyta, Viridiplantae, and Glaucophyta). Rhodoplasts and chloroplasts were later spread horizontally into other eukaryotic lineages through secondary endosymbiosis. Primary endosymbiosis is therefore a critical feature of eukaryotic evolution; however, it is difficult to study because of the long evolutionary time span that has passed since primary plastid origin. The filose amoeba *Paulinella chromatophora* is an exceptional species that contains two plastids, referred to as "chromatophores," that originated from a *Synechococcus*-like cyanobacterium. Photosynthetic *Paulinella* provides an ideal model to gain insights into the origin of photoautotrophy because its sister species are all heterotrophs that prey on cyanobacteria. Here, we review the evolutionary process that led to this second instance of primary endosymbiosis based on recent studies that include biodiversity surveys and plastid and nuclear genome data. Draft genome data from heterotrophic *Paulinella* using the single-cell genomics approach demonstrate two cases of horizontal gene transfer (HGT) from cyanobacteria, demonstrating that prey items are potential sources of foreign

H.S. Yoon (✉)
Department of Biological Sciences, Sungkyunkwan University, Suwon 440-746, South Korea
e-mail: hsyoon2011@skku.edu

E.C. Yang
Marine Ecosystem Research Division, Korea Institute of Ocean Science & Technology, Ansan 426-744, South Korea

H. Qiu
Department of Ecology, Evolution and Natural Resources, Rutgers University, New Brunswick, NJ 08901, USA

D. Bhattacharya
Department of Ecology, Evolution and Natural Resources, Institute of Marine and Coastal Science, Rutgers University, New Brunswick, NJ 08901, USA

W. Löffelhardt (ed.), *Endosymbiosis*, DOI 10.1007/978-3-7091-1303-5_8,

© Springer-Verlag Wien 2014

DNA in these taxa. Genome data from photosynthetic *Paulinella* provide evidence of massive gene loss from the chromatophore genome, endosymbiotic gene transfer (EGT) to the host nucleus, and the potential establishment of a plastid protein import system that relies on the secretory pathway in the amoeba. We also present recent data regarding postendosymbiotic speciation in photosynthetic *Paulinella* and lineage specific differential gene loss and EGT.

Contents

Origin of Primary Plastids 152
Paulinella Acquired Its Plastid from a *Synechococcus*-Like Cyanobacterium 154
Heterotrophic Sister Taxa of *Paulinella* Feed on Cyanobacteria 156
Postendosymbiotic Speciation 157
Plastid Genomes of *Paulinella* 159
Evidence of Endosymbiotic Gene Transfer 160
Conclusions 162
References 163

Origin of Primary Plastids

More than 300,000 land plant species dominate terrestrial habitats and their closest relatives, the green algae (collectively referred to as Viridiplantae), thrive in aquatic environments. The latter compete for natural resources with the red algae that comprise more than 6,000 species (http://www.algaebase.org). Including the relatively species-poor glaucophyte algae (i.e., 14 spp.), these three eukaryotic lineages contain two membrane-bounded photosynthetic organelles (plastids), a distinct feature that unites them as the Supergroup Plantae or Archaeplastida (Cavalier-Smith 1998; Adl et al. 2005; Price et al. 2012). Plantae play important ecosystem roles because they form the base of many food chains on our planet and also are of high interest to evolutionary biologists. The latter is because red and green algae are the source of the plastid via secondary endosymbiosis in a multitude of other taxa such as "chromalveolates" and euglenids.

The origin of the Plantae plastid is explained by primary endosymbiosis, whereby a heterotrophic protist engulfed and retained a cyanobacterium as a cytoplasmic organelle (Mereschkowsky 1905; Margulis 1970; Palmer 2003; Bhattacharya et al. 2004). A spate of past phylogenetic studies using nuclear loci has suggested Plantae polyphyly (e.g., Nozaki et al. 2007; Stiller 2007; Parfrey et al. 2010), implying multiple (primary or potentially secondary) endosymbiosis to explain plastid origin in these taxa. Other sources of genome information including a variety of plastid- and nuclear-encoded sequence data are however consistent with a single ancestor of the three primary plastid-containing groups (Rodriguez-Ezpeleta et al. 2005, 2007; Rodriguez-Ezpeleta and Philippe 2006; Burki et al. 2007, 2008; Hackett et al. 2007; Reyes-Prieto and Bhattacharya 2007a, b). Plantae monophyly is also supported by studies of landmark features such as the

composition and evolutionary history of protein translocons (Steiner et al. 2005; Weber et al. 2006; Gross and Bhattacharya 2009; Kalanon and McFadden 2008), origin of Calvin cycle enzymes (Reyes-Prieto and Bhattacharya 2007a, b), and perhaps most convincingly, the presence of dozens of Chlamydiae-derived genes from ancestral horizontal gene transfers (HGTs) that are shared by the Plantae lineages (Huang and Gogarten 2007; Moustafa et al. 2008). In agreement with these data, a rich novel gene repertoire (8,355 protein coding regions) from the completed genome of the unicellular red alga *Porphyridium purpureum* and partial genome data from the crustose red alga *Calliarthron tuberculosum* strongly supports the monophyly of the red and Viridiplantae lineages (Chan et al. 2011), and most recently analysis of complete genome data from the glaucophyte *Cyanophora paradoxa* (Price et al. 2012) now firmly supports monophyly of all three extant Plantae lineages. Therefore, it is highly likely that, after a single primary endosymbiosis between a nonphotosynthetic protist and a cyanobacterium, the first algal populations gave rise to the common ancestor of the red algae, glaucophyte algae, and the Viridiplantae. It is noteworthy that the plastids of Plantae have been horizontally spread into diverse eukaryotic groups via secondary endosymbiosis. A red algal plastid was transferred to the ancestor of "chromalveolates" (e.g., cryptomonads, haptophytes, heterokonts, and dinoflagellates), whereas green algal cells were independently captured in the ancestors of euglenids and chlorarachniophytes (Gibbs 1978, 1981; Cavalier-Smith 1999; McFadden 1999; Tengs et al. 2000; Yoon et al. 2002, 2005). Endosymbiosis has therefore been a driving force of eukaryote diversification.

Endosymbiosis research, however, has been hampered by the antiquity of Plantae plastid origin. During the 1.5 billion year evolutionary time span from the primary endosymbiosis [Yoon et al. 2004; see also Douzery et al. (2004)], cyanobacterial endosymbiont genomes (e.g., 4.66 Mbp in *Gloeobacter* PCC 7421) have been reduced to 100–150 kbp in extant organelles. Most of the genes have been lost outright or transferred to the host nuclear genome, where the molecular signatures of their history have been weakened or completely erased as a result of the evolution of novel functions. The rarity of endosymbiosis is largely due to the complex process required to refine and retool the host–plastid relationships including, endosymbiotic gene transfer (EGT), protein import from the host cytoplasm to the plastid, and the establishment of an effective metabolic connection between the partners, including of course the transport of fixed carbon from the nascent organelle to the host cytosol (Reyes-Prieto et al. 2007). The filose amoeba *Paulinella chromatophora* (Fig. 1a, b) provides an outstanding model to better understand the process of endosymbiosis because it is the only known case of an independent primary endosymbiosis between a heterotrophic protist, *Paulinella*, and a *Synechococcus*-like cyanobacterium. The "chromatophore" (here, used interchangeably with plastid) in these taxa was captured about 60 Ma and has features that indicate a "work in progress" with regard to plastid genome evolution (Marin et al. 2005; Yoon et al. 2006; Bhattacharya et al. 2007; Nowack et al. 2008, 2010; Reyes-Prieto et al. 2010; Nowack and Grossman 2012). Here, we review the outcome of past and more recent *Paulinella* research with a focus on the origin of the endosymbiont, postendosymbiotic speciation, chromatophore genome evolution, and HGT.

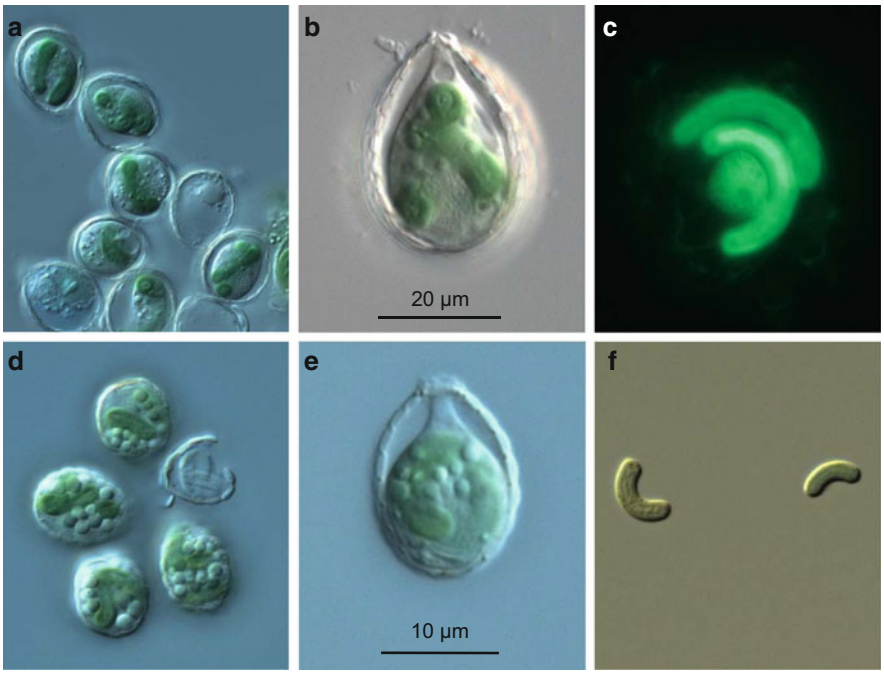


Fig. 1 Light micrograph images of *Paulinella chromatophora* CCAC 0185 (**a**–**c**) and *Paulinella* sp. FK01 (**d**–**f**). Images of two plastids stained by SYBR Green (**c**), and isolated plastids (**f**)

Paulinella Acquired Its Plastid from a *Synechococcus*-Like Cyanobacterium

The two blue–green chromatophores in *P. chromatophora* (Bhattacharya et al. 1995) were initially described by Lauterborn (1895) and retain key cyanobacterial features such as peptidoglycan, carboxysomes, and a similar gross morphology (Fig. 1c, f). In the original description, Lauterborn postulated the possible endosymbiotic origin of the cellular organelles with the statement, "... Cyanophyceae—that live with the rhizopod in an intimate symbiosis, or finally they are integral components, real organs of the rhizopod cell body" (Lauterborn 1895; Melkonian and Mollenhauer 2005). The chromatophores lie free in the cytoplasm, are not bounded by a vacuolar membrane, and the mature cell always contains two chromatophores, which divide after host cell division (Kies 1974; Kies and Kremer 1979; Johnson et al. 1988). Furthermore, chromatophores are readily degraded when isolated from the host cytosol (Fig. 1f) and cannot be cultured alone (Lauterborn 1895). Photosynthetic *Paulinella* (e.g., FK01 and CCAC0185 strains) have been maintained long term in culture without an external carbon source, suggesting true photoautotrophy (Nowack et al. 2008; Yoon et al. 2009).

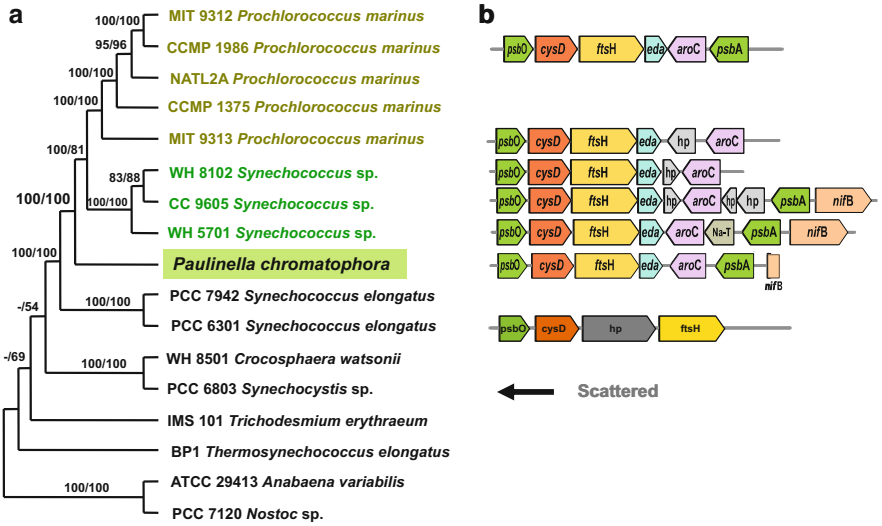


Fig. 2 (**a**) Maximum likelihood tree of concatenated plastid proteins (*fts*H, *psb*A, *psb*O, and *tuf*A). *Paulinella chromatophora* shows a sister group relationship to alpha-cyanobacteria including *Prochlorococcus–Synechococcus* strains. (**b**) Alignment of the 9.4 kbp plastid genome fragment from *Paulinella chromatophora* with homologous regions in closely related cyanobacteria (after Yoon et al. 2006)

Taken together, these features indicate a bona fide cellular organelle status for chromatophores.

Despite its vital importance to elucidating organelle evolution, *P. chromatophora* occurs rarely in nature and is challenging to maintain in culture (Kies 1974; Kies and Kremer 1979). However, the availability of two recently established culture strains of *P. chromatophora* CCAC 0185 (=M0880/a, Fig. 1a–c) and *Paulinella* FK01 (Fig. 1d–f) led to intensive research work with this taxon (Marin et al. 2005, 2007; Yoon et al. 2006, 2009; Nowack et al. 2008, 2010; Reyes-Prieto et al. 2010). Nuclear small subunit rDNA sequence data showed the amoeba host to be affiliated with the Euglyphida (Phylum Cercozoa; Supergroup Rhizaria), which comprises heterotrophic taxa except *P. chromatophora* (Bhattacharya et al. 1995). In contrast, plastid rDNA trees showed the *Paulinella* plastid to be most closely related to *Synechococcus* WH5701 within a larger clade of *Prochlorococcus* and *Synechococcus* species (Marin et al. 2005; Yoon et al. 2006). This cyanobacterial group, referred to as alpha-cyanobacteria, contains a unique RuBisCo Form 1A that is distinguishable from that found in other cyanobacteria and in plastids (Marin et al. 2007; Criscuolo and Gribaldo 2011). A multigene phylogeny using a concatenated protein dataset of *fts*H, *psb*A, *psb*O, and *tuf*A confirmed the chromatophore to have originated from a member of the alpha-cyanobacteria (Fig. 2). Gene synteny data also presented in this study also supported the alpha-cyanobacterial origin of the plastid. The alignment of the chromatophore genome fragment with homologous regions in alpha-cyanobacteria,

shows a strong conservation of gene order relative to *Synechococcus* sp. WH5701 reflecting a recent evolutionary history for plastid establishment in *Paulinella* (Yoon et al. 2006). Therefore, it appears to be clear that *Paulinella* acquired its plastid from a *Synechococcus*-like alpha-cyanobacterium.

Heterotrophic Sister Taxa of *Paulinella* Feed on Cyanobacteria

A total of nine heterotrophic *Paulinella* species have been reported to date. Among them, three species (*P. ovalis*, *P. intermedia*, and *P. indentata*) have long been known (Johnson et al. 1988; Vørs 1993; Hannah et al. 1996), whereas the other six marine species (*P. carsoni, P. agassizi, P. suzukii, P. lauterborni, P. multipora,* and *P. gigantica*) were recently described from three sand beach samples collected in British Columbia, Canada (Nicholls 2009). *Paulinella* species range from 5 to 47 μm in length. Heterotrophic species were reported from coastal and brackish waters as benthic and planktonic forms, whereas photosynthetic *Paulinella* thrives in small freshwater ponds. *P. indentata* cells were isolated from a freshwater pond near a brackish river in Adelaide, Australia and provide a possible scenario for the origin of freshwater species from their marine sisters (Yoon and Yang, unpublished).

It is noteworthy that *P. ovalis* ingests cyanobacteria in food vacuoles (Johnson et al. 1988). Therefore, it is reasonable to assume that the primary plastid endosymbiosis occurred in a heterotrophic ancestor of *P. chromatophora* that preyed on cyanobacteria, and at some point maintained a cyanobacterium that was free in the cytosol rather than being digested in a vacuole. As described below, aspects of this hypothesis recently gained some support in work from our lab (Bhattacharya et al. 2012). Because *P. ovalis*-like cells are too small (less than 5 μm) to be isolated by hand, the single-cell genomics method was applied to circumvent this problem. Single-cell genomics provides a powerful tool to generate draft genome data from cells captured in nature, with minimal chance for contamination (Stepanauskas and Sieracki 2007; Yoon et al. 2011). The procedure consists of three steps (1) single-cell sorting using flow cytometry, (2) whole genome amplification, and (3) taxonomic identification using conserved 18S rDNA sequence data (Heywood et al. 2011). Based on draft genomes of six *P. ovalis*-like cells isolated from Chesapeake Bay, USA, 34 genes were identified as being presumably derived from cyanobacterial prey. Among them, 12 are related specifically to marine *Prochlorococcus* and *Synechococcus*, the closest known relatives of the chromatophore. Most interestingly, two examples of cyanobacterial-derived horizontal gene transfer (HGT) were identified (i.e., DAP epimerase and leucyl-tRNA synthetase) in the nuclear DNA. The presence of three spliceosomal introns proves a nuclear origin of the *dap*F gene (for DAP epimerase) that shows a phylogenetic affinity to the alpha-cyanobacteria, whereas its two flanking genes are of eukaryotic origin

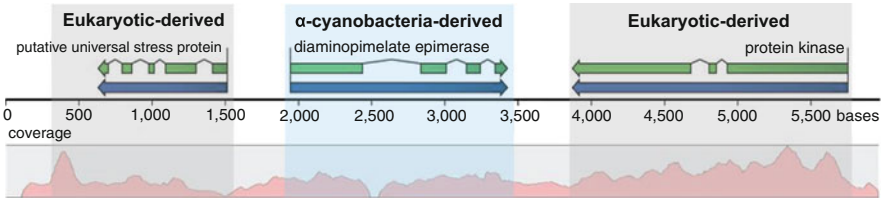


Fig. 3 Genomic region of the heterotrophic *Paulinella ovalis*-like cell that shows an example of HGT from an alpha-cyanobacterial source. Intron distribution and coverage of *P. ovalis*-like genome contig ConsensusPlus1618 that encodes three proteins. DAP epimerase from *P. ovalis*-like cells shows a close phylogenetic relationship to homologs from alpha-cyanobacteria and from the plastid genome of *P. chromatophora* CCAC 0185, whereas the two flanking genes show an affinity to eukaryote sequences (Bhattacharya et al. 2012)

(Fig. 3). This work provides the first evidence of a link between feeding behavior in wild-caught cells, HGT, and plastid primary endosymbiosis in the monophyletic *Paulinella* clade.

Postendosymbiotic Speciation

There is little doubt that speciation was one of the most important evolutionary processes that Darwin addressed (Darwin 1859). Given the 60 Ma minimum age of the *Paulinella* plastid, it is therefore of high interest to find other photosynthetic *Paulinella* species to facilitate an in-depth analysis of postendosymbiotic genome evolution in distinct lineages that share a common ancestral endosymbiont. In fact, since Lauterborn's first description from Germany, photosynthetic *Paulinella* (i.e., *P. chromatophora*) has been reported from around the world, including sites in Switzerland (Pendard 1905), the UK (Brown 1915), and the USA (Kepner 1905; Lackey 1936). However, these reports were simple statements of occurrence without any detailed description or deposition of vouchers. All published morphological and ultrastructural studies thus far stem from samples collected in Germany (Kies 1974; Kies and Kremer 1979). The recent molecular phylogenetic and genomic studies also relied on the culture CCAC 0185 (=M0880/a) that was isolated in Germany (Bhattacharya et al. 1995; Marin et al. 2005, 2007; Yoon et al. 2006; Nowack et al. 2008). This scant history of collection may reflect the rarity of *P. chromatophora* in nature, or simply an absence of a comprehensive effort to isolate novel taxa.

A new culture strain FK01 was isolated from a freshwater site in Japan (Yoon et al. 2009). The morphological characters of *Paulinella* FK01 clearly distinguished it from the original descriptions as well as the existing culture strain CCAC 0185 in cell size (17 × 11 versus 27 × 20 μm), scales per each column (10–11 versus 12–14), and number of oral scales (5 versus 3). More distinctive, multiple fine pores

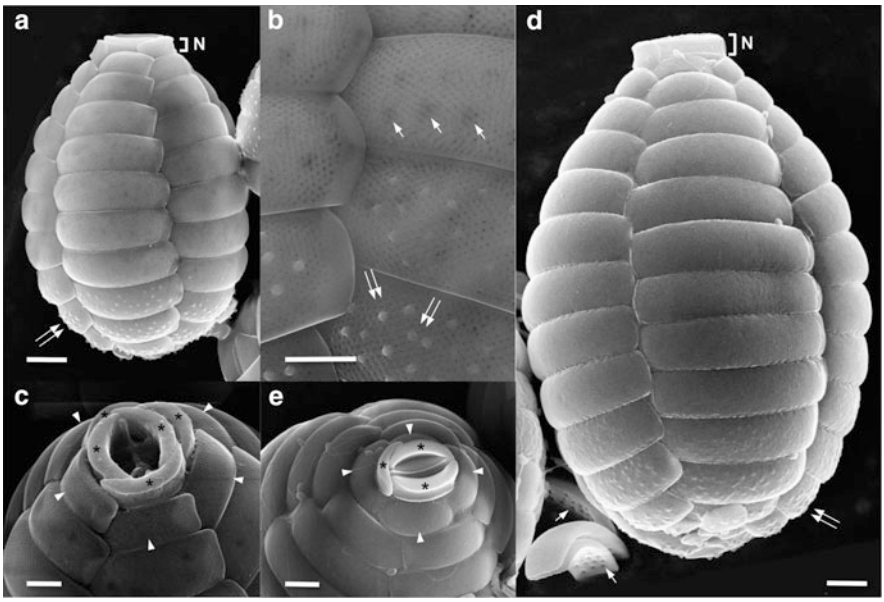


Fig. 4 SEM images of photosynthetic *Paulinella* sp. FK01 (**a**–**c**) and *P. chromatophora* CCAC 0185 (**d**, **e**). FK01 is smaller in cell size than CCAC 0185. Distinctive, multiple fine pores are present on the surface of scales of FK01 (**b**). Five oral scales occurred in FK01 (**c**), whereas only three were found in CCAC 0185 (**e**) (adapted from Yoon et al. 2009)

occurred on the external surface of scales in *Paulinella* FK01 (see Fig. 4). Papillae on the four posterior scales were more distinctive than those found in CCAC 0185. In addition to the obvious morphological differences, molecular data demonstrate sequence divergence between CCAC 0185 and FK01. Phylogenies from nuclear 18S rDNA and actin data support the common origin of the host ancestor, which was sister to other euglyphids within the Cercozoa, whereas the plastid 16S + 23S rDNA tree show the plastid to be of alpha-cyanobacterial origin (Yoon et al. 2006, 2009; Nowack et al. 2008). Pairwise analysis of synonymous (K_s) and nonsynonymous (K_a) substitution rates between the *Paulinella* actin coding regions were 1.0023 and 0.0181, respectively ($K_a/K_s = 0.0181$). This ratio was comparable to actin sequence differences between two green algal *Ostreococcus* species (i.e., *O. tauri* vs. *O. lucimarinus*; $K_a/K_s = 0.0068$) and between yeasts (*Saccharomyces cerevisiae* vs. *Pichia stipitis*, $K_a/K_s = 0.0283$). These results suggest that CCAC 0185 and FK01 are significantly diverged from each other and likely constitute distinct species [i.e., *P. microporus* sp. nov., Yoon et al. (in preparation)]. These data indicate a single origin of the chromatophore, after which novel species diverged within the photosynthetic *Paulinella* lineage (Fig. 5).

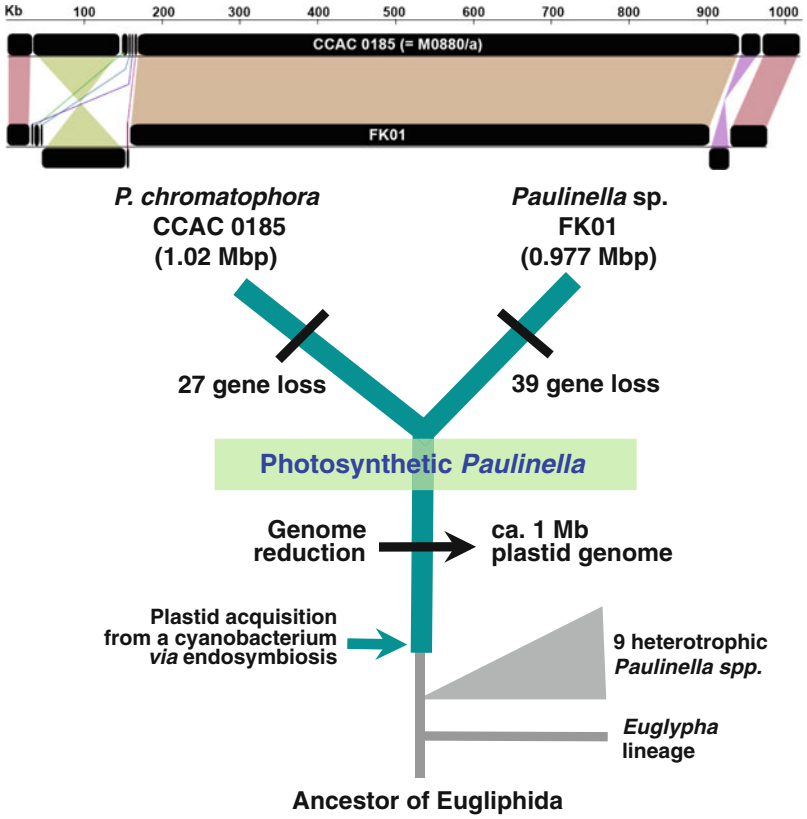


Fig. 5 Major events in the evolution of the genus *Paulinella* and an alignment of the plastid genome from *Paulinella* CCAC 0185 and FK01. After primary endosymbiosis between a heterotrophic *Paulinella* cell and a *Synechococcus*-like cyanobacterium, massive chromatophore gene loss has occurred. Ancestral photosynthetic *Paulinella* diverged into two species: *P. chromatophora* CCAC 0185 and *Paulinella* sp. FK01. Each species shows differential plastid gene loss. The two plastid genomes show overall conservation of gene order with five genome inversions that are indicated with the *diagonal lines*

Plastid Genomes of *Paulinella*

Two complete chromatophore genome sequences have been published thus far, from *P. chromatophora* CCAC 0185 (Nowack et al. 2008) and from *Paulinella* FK01 (Reyes-Prieto et al. 2010). The CCAC 0185 work followed up on a partial chromatophore genome sequence from this strain that was published by Yoon et al. (2006). The plastid genome of *Paulinella* CCAC 0185 encodes 867 protein-coding genes, 42 tRNAs, 2 rRNA clusters, and 223 hypothetical proteins on a 1.02 Mbp circular molecule. Compared to *Synechococcus* sp. WH5701 (3.04 Mbp

with 3,346 protein-coding genes), *Paulinella* CCAC 0185 retains only 26 % of the original gene content. Furthermore, it contains 11 putative pseudogenes that suggest genome reduction is ongoing. Compared to the gene family-rich cyanobacterial genome, *Paulinella* maintains only single-copy genes except for *psb*A, *psb*D, and the rDNA gene clusters. *Paulinella* provides direct evidence of gene loss in redundant paralogs (Mendonca et al. 2011; Qiu et al. 2012). An interesting question is whether significant genome reduction in the *Paulinella* chromatophore was accompanied by loss of essential genes and pathways. Nowack and colleagues (2008) found that all the genes were missing for essential pathways involved in amino acid and cofactor synthesis, as well, there were single genes missing from other biosynthetic pathways. Based on these results, it was postulated that the chromatophore is completely dependent on the host amoeba for its survival. Presumably, some missing genes from biosynthetic pathways in the chromatophore have been replaced by preexisting host genes, or alternatively, the chromatophore genes were not lost outright but relocated to the nuclear genome. Either scenario requires import of cytosolic proteins to the chromatophore via a targeting system (see below). Taken together, genome data clearly demonstrate that the chromatophore is a true cellular organelle, despite earlier controversy about its status as a bona fide organelle or an endosymbiont (Bhattacharya and Archibald 2006; Theissen and Martin 2006; Bodyl et al. 2007, 2010).

Compared to the CCAC 0185 plastid, the second sequenced plastid genome from *Paulinella* FK01 is smaller in size (977 kbp) and encodes 841 predicted proteins and 48 stable RNAs. Inspection of the genomes reveals overall conservation of gene order with five inversions (Reyes-Prieto et al. 2010; see Fig. 5). It is interesting that gene-by-gene comparisons of the FK01 and CCAC 0185 plastid genomes revealed 27 genes encoded in FK01 that are absent from CCAC 0185, whereas 39 genes in CCAC 0185 are absent from FK01 (see Fig. 5). These 66 genes all have an alpha-cyanobacterial origin and were likely present in the common ancestor of these photosynthetic *Paulinella* and provide examples of lineage-specific gene loss. Inspection of 681 DNA alignments of protein-encoding genes shows that the vast majority of genes were under purifying selection (i.e., K_a/K_s ratios $\ll 1$). Taken together, gene order and differential gene loss in plastid genomes from *Paulinella* provide for the first time clear examples of the nature and tempo of postendosymbiotic genome evolution for an organelle of recent origin.

Evidence of Endosymbiotic Gene Transfer

Endosymbiotic gene transfer (EGT) is associated with organelle origin and contributes significantly to nuclear genome evolution. There is marked plastid genome reduction in *Paulinella* species, which have highly reduced genome sizes (ca. 1 Mbp) from an ancestral ca. 3 Mbp in free-living cyanobacteria of the putative donor *Synechococcus* clade (Nowack et al. 2008; Reyes-Prieto et al. 2010). Therefore, about 2 Mbp of cyanobacterial genome sequence was either lost outright or

relocated to the host nucleus via primary EGT. This massive amount of genome reduction could be accompanied by EGT of essential genes into the host nucleus to maintain a functional but dependent endosymbiont. This necessitates the evolution of an efficient protein import system to deliver chromatophore-destined proteins that are translated in the host cytosol back to the endosymbiont. It is therefore important to better understand how EGT has contributed to host genome evolution to gain insights into organellogenesis (e.g., Gross and Bhattacharya 2009).

Two instances of EGT for *psa*E and *psa*I were reported in *Paulinella* FK01 (Nakayama and Ishida 2009; Reyes-Prieto et al. 2010). In particular, cyanobacterial *psa*I that encodes subunit VIII of photosystem I (PSI) has been silenced by two nonsense mutations in the *Paulinella* FK01 plastid genome, whereas an intact copy with a 198 bp spliceosomal intron exists in the nucleus of the host. Plastid encoded *psa*I was therefore likely pseudogenized after activation of the transferred nuclear copy. In contrast, CCAC 0185 retains *psa*I in the plastid genome suggesting that lineage-specific EGT of *psa*I occurred after the divergence of these two species. Recently, Nowack et al. (2010) reported 32 examples (i.e., *psa*K, *ycf*34, *cso*S4A, *psb*N, CP12, multicopies of *hli*, and hypothetical proteins) of EGT from an analysis of 32,012 ESTs from *P. chromatophora* CCAC 0185. They postulated that 0.3–0.8 % of nuclear genes in the amoeba were obtained via EGT, significantly lower than the 11–14 % postulated for algal members of the Plantae (Reyes-Prieto et al. 2006). The majority of EGT candidates were involved in photosynthesis and electron transport with a regulatory function. EGT candidates showed a GC content of 52.5 %, which is close to that of nuclear genes (49.6 %) but clearly different from plastid-encoded genes (40.4 %). They found a recent (i.e., postspeciation) example of intron insertion in different regions of *psa*E that distinguish *P. chromatophora* CCAC 0185 and *Paulinella* FK01. Furthermore, these authors reported two expressed genes (*hli* and *psa*K) that are plastid encoded in *Paulinella* FK01 but absent in the *P. chromatophora* CCAC 0185 organelle, also suggesting differential EGT postspeciation.

Bodyl and colleagues conducted a series of bioinformatic analyses to search for the protein trafficking system between the *Paulinella* host nucleus and chromatophore (Bodyl et al. 2007, 2010; Mackiewicz et al. 2012). They found homologs of Tic21, Tic32, and Toc12 from the CCAC 0185 plastid genome, which lacked the Omp85/Toc75 and Tic20 genes. They postulated that these missing genes had been transferred to the *Paulinella* nucleus where after translation in the cytosol the encoded proteins were imported and integrated into the endosymbiont membranes (Bodyl et al. 2010). Following an analysis of 10 EGT-derived gene products from CCAC 0185, they identified potential signal peptides in five proteins and a putative transit peptide from one protein that might be involved in host ER membrane-mediated or endosymbiont envelope-mediated import, respectively. Based on the low molecular weight and nearly neutral charge of EGT-derived proteins, they predicted that these proteins pass freely through the peptidoglycan wall between the chromatophore outer and inner membranes (Mackiewicz et al. 2012).

The bioinformatic predictions described above and the resulting models are interesting exercises, but they lack explanatory power without cell biological data.

The first such data were recently produced by Nowack and Grossman (2012) and provide the first direct evidence for a functional plastid import system in photosynthetic *Paulinella* (and for that matter, in any host–bacterial endosymbiont system, excluding plastids and mitochondria). These authors used *psa*E, *psa*K1, and *psa*K2 as models for their work. Using immunogold electron microscopy (EM) with labeled α-psaE antibodies, they showed nuclear-encoded psaE protein to be clearly localized to the chromatophore and also to be associated with the Golgi apparatus. They also showed that these PSI subunits (*psa*E, *psa*K1 and *psa*K2) are synthesized in the cytoplasm in association with 80S ribosomes, whereas they are assembled into PSI complexes with the remaining plastid-encoded subunits within the chromatophore. These data suggest that the host secretory pathway is sufficient to support chromatophore protein import. This sort of primitive import mechanism has been previously postulated to be an early step in translocon evolution in other algae and plants (e.g., Bhattacharya et al. 2007; Reyes-Prieto et al. 2007). Taken together, EGT in *Paulinella* is an ongoing process, and consequently the protein import machinery is likely also evolving to allow more efficient trafficking of nuclear-encoded plastid-targeted proteins to these organelles.

Conclusions

Morphological, molecular, and genome data strongly support the idea that the phagotrophic *Paulinella* lineage acquired its plastid from a *Synechococcus*-like cyanobacterium. Given the presence of alpha-cyanobacterium-derived genes in the nuclear genome of heterotrophic *Paulinella*, it is highly likely that HGT occurred in heterotrophic *Paulinella*. It is still unknown whether cyanobacterium-derived genes (i.e., DAP epimerase and leucyl-tRNA synthetase) are expressed and maintain the original functions in the cytosol of heterotrophic *P. ovalis*-like cells; however, it is conceivable that these types of cyanobacterium-derived HGTs could have supported the primary endosymbiosis in the photosynthetic *Paulinella* lineage. Once successful, organellogenesis led to two-thirds of the endosymbiont genome, mostly redundant genes, being lost prior to speciation. Two photosynthetic *Paulinella* species have been characterized with four more new lineages under investigation (Yoon et al., in preparation). Less than 100 genes are likely to have been transferred from the endosymbiont to the host nucleus including many essential genes for photosynthetic function. The plastid protein import machinery is likely to still be "under development" and is not yet fully understood. Upcoming complete genome data from photosynthetic *Paulinella* species promises to further unravel how a free-living cyanobacterium becomes an obligate cellular organelle.

References

Adl SM, Simpson AG, Farmer MA, Andersen RA, Anderson OR, Barta JR, Bowser SS, Brugerolle G, Fensome RA, Fredericq S, James TY, Karpov S, Kugrens P, Krug J, Lane CE, Lewis LA, Lodge J, Lynn DH, Mann DG, McCourt RM, Mendoza L, Moestrup O, Mozley-Standridge SE, Nerad TA, Shearer CA, Smirnov AV, Spiegel FW, Taylor MF (2005) The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. J Eukaryot Microbiol 52:399–451

Bhattacharya D, Archibald JM (2006) Response to Theissen and Martin. Curr Biol 16: R1017–R1018

Bhattacharya D, Archibald JM, Weber AP, Reyes-Prieto A (2007) How do endosymbionts become organelles? Understanding early events in plastid evolution. Bioessays 29:1239–1246

Bhattacharya D, Helmchen T, Melkonian M (1995) Molecular evolutionary analyses of nuclear-encoded small subunit ribosomal RNA identify an independent rhizopod lineage containing the Euglyphina and the Chlorarachniophyta. J Eukaryot Microbiol 42:65–69

Bhattacharya D, Price DC, Yoon HS, Yang EC, Poulton NJ, Andersen RA, Das SP (2012) Single cell genome analysis supports a link between phagotrophy and primary plastid endosymbiosis. Sci Rep 2:356

Bhattacharya D, Yoon HS, Hackett JD (2004) Photosynthetic eukaryotes unite: endosymbiosis connects the dots. BioEssays 26:50–60

Bodyl A, Mackiewicz P, Stiller JW (2007) The intracellular cyanobacteria of *Paulinella chromatophora*: endosymbionts or organelles? Trends Microbiol 15:295–296

Bodyl A, Mackiewicz P, Stiller JW (2010) Comparative genomic studies suggest that the cyanobacterial endosymbionts of the amoeba *Paulinella chromatophora* possess an import apparatus for nuclear-encoded proteins. Plant Biol (Stuttg) 12:639–649

Brown JM (1915) On the occurrence of *Paulinella chromatophora* in Britain. Naturalist 157–159

Burki F, Shalchian-Tabrizi K, Minge M, Skjaeveland A, Nikolaev SI, Jakobsen KS, Pawlowski J (2007) Phylogenomics reshuffles the eukaryotic supergroups. PLoS One 2:e790

Burki F, Shalchian-Tabrizi K, Pawlowski J (2008) Phylogenomics reveals a new 'megagroup' including most photosynthetic eukaryotes. Biol Lett 4:366–369

Cavalier-Smith T (1998) A revised six-kingdom system of life. Biol Rev 73:203–266

Cavalier-Smith T (1999) Principles of protein and lipid targeting in secondary symbiogenesis: euglenoid, dinoflagellate, and sporozoan plastid origins and the Eukaryote family tree. J Eukaryot Microbiol 46:347–366

Chan CX, Yang EC, Banerjee T, Yoon HS, Martone PT, Estevez JM, Bhattacharya D (2011) Red and green algal monophyly and extensive gene sharing found in a rich repertoire of red algal genes. Curr Biol 21:328–333

Criscuolo A, Gribaldo S (2011) Large-scale phylogenomic analyses indicate a deep origin of primary plastids within cyanobacteria. Mol Biol Evol 28:3019–3032

Darwin C (1859) On the origin of species by means of natural selection or the preservation of favoured races in the struggle for life. Reprinted. Gramercy, London

Douzery EJ, Snell EA, Bapteste E, Delsuc F, Philippe H (2004) The timing of eukaryotic evolution: does a relaxed molecular clock reconcile proteins and fossils? Proc Natl Acad Sci USA 101:15386–15391

Gibbs SP (1978) The chloroplasts of *Euglena* may have evolved from symbiotic green algae. Can J Bot 56:2883–2889

Gibbs SP (1981) The chloroplasts of some algal groups may have evolved from endosymbiotic eukaryotic algae. Ann N Y Acad Sci 361:193–208

Gross J, Bhattacharya D (2009) Revaluating the evolution of the Toc and Tic protein translocons. Trends Plant Sci 14:13–20

Hackett JD, Yoon HS, Li S, Reyes-Prieto A, Rummele SE, Bhattacharya D (2007) Phylogenomic analysis supports the monophyly of cryptophytes and haptophytes and the association of rhizaria with chromalveolates. Mol Biol Evol 24:1702–1713

Hannah F, Rogerson A, Anderson OR (1996) A description of *Paulinella indentata* n. sp. (Filosea: Euglyphina) from subtidal coastal benthic sediments. J Eukaryot Microbiol 43:1–4

Heywood JL, Sieracki ME, Bellows W, Poulton NJ, Stepanauskas R (2011) Capturing diversity of marine heterotrophic protists: one cell at a time. ISME J 5:674–684

Huang J, Gogarten JP (2007) Did an ancient chlamydial endosymbiosis facilitate the establishment of primary plastids? Genome Biol 8:R99

Johnson PW, Hargraves PE, Sieburth JM (1988) Ultrastructure and ecology of *Calycomonas ovalis* Wulff, 1919, (Chrysophyceae) and its redescription as a testate rhizopod, *Paulinella ovalis* n. comb. (Filosea: Euglyphina). J Protozool 35:618–626

Kalanon M, McFadden GI (2008) The chloroplast protein translocation complexes of *Chlamydomonas reinhardtii*: a bioinformatic comparison of Toc and Tic components in plants, green algae and red algae. Genetics 179:95–112

Kepner WA (1905) *Paulinella chromatophora*. Biol Bull 9:128–129

Kies L (1974) Elektronenmikroskopische untersuchungen an *Paulinella chromatophora* Lauterborn, einer thekamobe mit blau-grunen endosymbionten (Cyanellen). Protoplasma 80:69–89

Kies L, Kremer BP (1979) Function of cyanelles in the thecamoeba *Paulinella chromatophora*. Naturwissenschaften 66:578

Lackey JB (1936) Some fresh water protozoa with blue chromatophores. Biol Bull 71:492–497

Lauterborn R (1895) Protozoenstudien II. *Paulinella chromatophora* nov. gen., nov. spec., ein beschalter Rhizopode des Sußwassers mit blaugrunen chromatophorenartigen Einschlussen. Z Wiss Zool 59:537–544

Mackiewicz P, Bodyl A, Gagat P (2012) Possible import routes of proteins into the cyanobacterial endosymbionts/plastids of *Paulinella chromatophora*. Theory Biosci 131:1–18

Margulis L (1970) Origin of eukaryotic cells. Yale University Press, New Haven, CT

Marin B, Nowack EC, Glockner G, Melkonian M (2007) The ancestor of the *Paulinella* chromatophore obtained a carboxysomal operon by horizontal gene transfer from a *Nitrococcus*-like gamma-proteobacterium. BMC Evol Biol 7:85

Marin B, Nowack EC, Melkonian M (2005) A plastid in the making: evidence for a second primary endosymbiosis. Protist 156:425–432

McFadden GI (1999) Endosymbiosis and evolution of the plant cell. Curr Opin Plant Biol 2:513–519

Melkonian M, Mollenhauer D (2005) Robert Lauterborn (1869–1952) and his *Paulinella chromatophora*. Protist 156:253–262

Mendonca AG, Alves RJ, Pereira-Leal JB (2011) Loss of genetic redundancy in reductive genome evolution. PLoS Comput Biol 7:e1001082

Mereschkowsky C (1905) Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. Biologisches Centralblatt 25:593–604

Moustafa A, Reyes-Prieto A, Bhattacharya D (2008) Chlamydiae has contributed at least 55 genes to Plantae with predominantly plastid functions. PLoS One 3:e2205

Nakayama T, Ishida K (2009) Another acquisition of a primary photosynthetic organelle is underway in *Paulinella chromatophora*. Curr Biol 19:R284–R285

Nicholls K (2009) Six new marine species of the genus *Paulinella* (Rhizopoda: Filosea, or Rhizaria: Cercozoa). J Mar Biol Assoc UK 89:1415–1425

Nowack EC, Melkonian M, Glockner G (2008) Chromatophore genome sequence of *Paulinella* sheds light on acquisition of photosynthesis by eukaryotes. Curr Biol 18:410–418

Nowack EC, Vogel H, Groth M, Grossman AR, Melkonian M, Glockner G (2010) Endosymbiotic gene transfer and transcriptional regulation of transferred genes in *Paulinella chromatophora*. Mol Biol Evol 28:407–422

Nowack EC, Grossman AR (2012) Trafficking of protein into the recently established photosynthetic oreganelles of *Paulinella chromatophora*. Proc Natl Acad Sci USA 109:5340–5345

Nozaki H, Iseki M, Hasegawa M, Misawa K, Nakada T, Sasaki N, Watanabe M (2007) Phylogeny of primary photosynthetic eukaryotes as deduced from slowly evolving nuclear genes. Mol Biol Evol 24:1592–1595

Palmer JD (2003) The symbiotic birth and spread of plastids: how many times and whodunit? J Phycol 39:4–12

Parfrey LW, Grant J, Tekle YI, Lasek-Nesselquist E, Morrison HG, Sogin ML, Patterson DJ, Katz LA (2010) Broadly sampled multigene analyses yield a well-resolved eukaryotic tree of life. Syst Biol 59:518–533

Pendard E (1905) Notes sur quelques sarcodines. Int Rev Suisse Zool 13:603–610

Price DC, Chan CX, Yoon HS, Yang EC, Qiu H, Weber APM, Schwacke R, Gross J, Blouin NA, Lane C, Reyes-Prieto A, Durnford DG, Neilson JAD, Lang BF, Burger G, Steiner JM, Löffelhardt W, Meuser JE, Posewitz MC, Ball S, Arias MC, Henrissat B, Coutinho PM, Rensing SA, Symeonidi A, Doddapaneni H, Green BR, Rajah VD, Boore J, Bhattacharya D (2012) *Cyanophora paradoxa* genome elucidates origin of photosynthesis in algae and plants. Science 335:843–847

Qiu H, Yang EC, Bhattacharya D, Yoon HS (2012) Ancient gene paralogy may mislead inference of plastid phylogeny. Mol Biol Evol 29:3333–3343

Reyes-Prieto A, Bhattacharya D (2007a) Phylogeny of Calvin cycle enzymes supports Plantae monophyly. Mol Phylogenet Evol 45:384–391

Reyes-Prieto A, Bhattacharya D (2007b) Phylogeny of nuclear-encoded plastid-targeted proteins supports an early divergence of glaucophytes within Plantae. Mol Biol Evol 24:2358–2361

Reyes-Prieto A, Hackett JD, Soares MB, Bonaldo MF, Bhattacharya D (2006) Cyanobacterial contribution to algal nuclear genomes is primarily limited to plastid functions. Curr Biol 16:2320–2325

Reyes-Prieto A, Weber AP, Bhattacharya D (2007) The origin and establishment of the plastid in algae and plants. Annu Rev Genet 41:147–168

Reyes-Prieto A, Yoon HS, Moustafa A, Yang EC, Andersen RA, Boo SM, Nakayama T, Ishida K, Bhattacharya D (2010) Differential gene retention in plastids of common recent origin. Mol Biol Evol 27:1530–1537

Rodriguez-Ezpeleta N, Brinkmann H, Burey SC, Roure B, Burger G, Loffelhardt W, Bohnert HJ, Philippe H, Lang BF (2005) Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes. Curr Biol 15:1325–1330

Rodriguez-Ezpeleta N, Brinkmann H, Burger G, Roger AJ, Gray MW, Philippe H, Lang BF (2007) Toward resolving the eukaryotic tree: the phylogenetic positions of jakobids and cercozoans. Curr Biol 17:1420–1425

Rodriguez-Ezpeleta N, Philippe H (2006) Plastid origin: replaying the tape. Curr Biol 16:R53–R56

Stepanauskas R, Sieracki ME (2007) Matching phylogeny and metabolism in the uncultured marine bacteria, one cell at a time. Proc Natl Acad Sci USA 104:9052–9057

Steiner JM, Yusa F, Pompe JA, Löffelhardt W (2005) Homologous protein import machineries in chloroplasts and cyanelles. Plant J 44:646–652

Stiller JW (2007) Plastid endosymbiosis, genome evolution and the origin of green plants. Trends Plant Sci 12:391–396

Tengs T, Dahlberg OJ, Shalchian-Tabrizi K, Klaveness D, Rudi K, Delwiche CF, Jakobsen KS (2000) Phylogenetic analyses indicate that the 19′Hexanoyloxy-fucoxanthin-containing dinoflagellates have tertiary plastids of haptophyte origin. Mol Biol Evol 17:718–729

Theissen U, Martin W (2006) The difference between organelles and endosymbionts. Curr Biol 16:R1016–R1017

Vørs N (1993) Marine heterotrophic amoebae, flagellates and Heliozoa from Belize (Central America) and Tenerife (Canary Islands), with descriptions of new species, *Luffisphaera bulbochaete* n. sp., *L. Longihastis* n. sp., *L. turriformis* n. sp. and *Paulinella intermedia* n. sp. J Eukaryot Microbiol 40:272–287

Weber AP, Linka M, Bhattacharya D (2006) Single, ancient origin of a plastid metabolite translocator family in Plantae from an endomembrane-derived ancestor. Eukaryot Cell 5:609–612

Yoon HS, Hackett JD, Ciniglia C, Pinto G, Bhattacharya D (2004) A molecular timeline for the origin of photosynthetic eukaryotes. Mol Biol Evol 21:809–818

Yoon HS, Hackett JD, Pinto G, Bhattacharya D (2002) The single, ancient origin of chromist plastids. Proc Natl Acad Sci USA 99:15507–15512

Yoon HS, Hackett JD, Van Dolah FM, Nosenko T, Lidie KL, Bhattacharya D (2005) Tertiary endosymbiosis driven genome evolution in dinoflagellate algae. Mol Biol Evol 22:1299–1308

Yoon HS, Nakayama T, Reyes-Prieto A, Andersen RA, Boo SM, Ishida KI, Bhattacharya D (2009) A single origin of the photosynthetic organelle in different *Paulinella* lineages. BMC Evol Biol 9:98

Yoon HS, Price DC, Stepanauskas R, Rajah VD, Sieracki ME, Wilson WH, Yang EC, Duffy S, Bhattacharya D (2011) Single-cell genomics reveals organismal interactions in uncultivated marine protists. Science 332:714–717

Yoon HS, Reyes-Prieto A, Melkonian M, Bhattacharya D (2006) Minimal plastid genome evolution in the *Paulinella* endosymbiont. Curr Biol 16:R670–R672

Rhopalodia gibba: The First Steps in the Birth of a Novel Organelle?

Susann Adler, Eike M. Trapp, Christine Dede, Uwe G. Maier, and Stefan Zauner

Abstract The diatom *Rhopalodia gibba* harbours unusual cell inclusions termed spheroid bodies. Those are separated from the host cytoplasm by an additional membrane, vertically transmitted to the next generation and provide nitrogen autonomy to their host cell. Morphological observations and phylogenetic analyses revealed a cyanobacterial origin of these obligate endosymbionts. Phylogenetic data and fossil records suggest that the origin of this endosymbiosis dates back to late Eocene to Miocene (~25 Ma). Genomic analyses support this determination, as the genomic changes in the spheroid body suggest that the endosymbiosis is in a relatively early stage.

Contents

Introduction 168
The Pennate Diatom *Rhopalodia gibba* 168
The Origin of the Spheroid Bodies 170
Localisation of the Nitrogen Fixing Activity 171
Initial Characterisation of the Spheroid Body Genome 171
Genome Size 172
Pseudogenes 173
Mobile Elements 174
G/C Content 174
DNA Repair 174

S. Adler • E.M. Trapp • C. Dede • S. Zauner (✉)
Laboratorium für Zellbiologie, Philipps-Universität Marburg, Karl-von-Frisch Str. 8, 35032 Marburg, Germany
e-mail: zauner@biologie.uni-marburg.de

U.G. Maier
Laboratorium für Zellbiologie, Philipps-Universität Marburg, Karl-von-Frisch Str. 8, 35032 Marburg, Germany

LOEWE-Zentrum für Synthetische Mikrobiologie (SynMikro), Hans-Meerwein-Strasse, 35032 Marburg, Germany

W. Löffelhardt (ed.), *Endosymbiosis*, DOI 10.1007/978-3-7091-1303-5_9,
© Springer-Verlag Wien 2014

The UCYN-A/Spheroid Body Relations 175
From Free-Living Cyanobacteria to Endosymbionts to *Paulinella chromatophora* to Plastids? 176
References 176

Introduction

Diatoms, a group of the phylum heterokontophyta, have fascinated biologists since their discovery by light microscopy probably in the year 1703 [reviewed in Round et al. (1990)]. Diatoms are unicellular, sometimes colonial algae that can be found in almost every aquatic and even some terrestrial habitats. They either live as benthic forms or as marine phytoplankton. According to their quantitative dominance and phototrophic lifestyle, diatoms contribute a significant part of oxygen evolution and CO_2 fixation in the atmosphere and hold a central position as primary producers in the global ecosystem (Hoek et al. 1993; Falkowski et al. 2004). About 100,000 existing species are divided into three classes: Coscinodiscophyceae (centric diatoms), Fragilariophyceae (pennate diatoms without a raphe) and Bacillariophyceae (pennate diatoms with a raphe) (Round et al. 1990).

The group of diatoms evolved via secondary endosymbiosis. Here, a photoautotrophic red alga was captured by a phagotrophic, probably phototrophic cell (Moustafa et al. 2009). Co-evolution of this cellular merger led to elimination of many compartments such as the nucleus of the symbiont. A so-called complex plastid is maintained from the secondary endosymbiont, which is surrounded by four membranes (Hempel et al. 2007; Bolte et al. 2009).

The Pennate Diatom Rhopalodia gibba

R. gibba (Bacillariophyceae, Fig. 1) belongs to the family Rhopalodiaceae (Round et al. 1990). As other members of this family, the diatoms harbour unusual cell inclusions (Geitler 1977), first described by Pfitzer (1869) and termed spheroid bodies. Three decades later, they were depicted as pyrenoids (Klebahn 1896). However, by use of electron microscopy, Drum and Pankratz (1965) disproved the latter description. They reported spheroid bodies; 4–6 μm wide and 5–7 μm long ovoid structures, as additional cell structures that are located within the cytoplasm separated from the host by one membrane (Fig. 2). The authors also described internal membranes that resemble thylakoids and DNA staining, and thereby assumed a symbiotic association between a cyanobacterium and *R. gibba* (Drum and Pankratz 1965).

Even though Drum and Pankratz (1965) speculated about the basic principle of this association, it took 15 years until Floener and Bothe were able to demonstrate the ability of *R. gibba* to fix molecular nitrogen (Floener and Bothe 1980). By use of

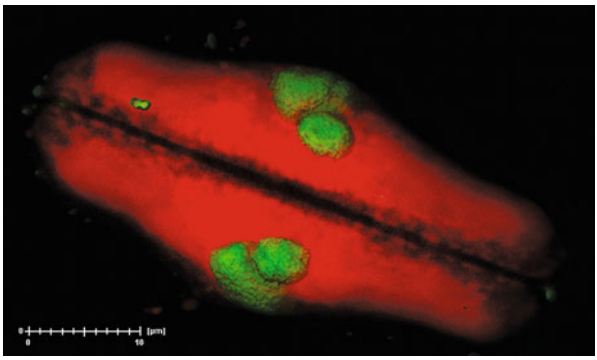


Fig. 1 Microscopic overview of *Rhopalodia gibba*. Cells (after division) were stained with SYBR Green I and examined by confocal microscopy. The chlorophyll autofluorescence of diatom plastid is shown in *red* while the DNA staining results in *green*. The two smaller DNA containing structures represent the nuclei of both daughter cells. The *larger green* structures correspond to the spheroid bodies of *Rhopalodia gibba*

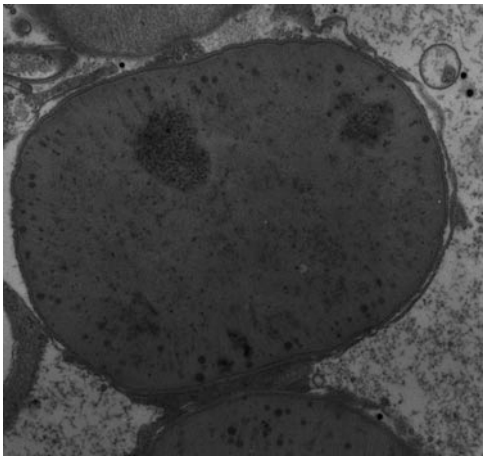


Fig. 2 Detailed illustration of the spheroid body. Spheroid bodies are separated from the cytosol by one host membrane and retain both cyanobacterial envelope membranes. The electron microscopic observation indicated internal membranes that resemble thylakoids

acetylene reduction assays, they identified the nitrogenase activity in the diatom and determined its light dependency (Floener and Bothe 1980). Interestingly, the addition of combined nitrogen in the form of nitrate or ammonia slightly decreased the nitrogenase activity but did not completely abolish nitrogen fixation as it would be expected when compared to cyanobacteria (Floener and Bothe 1980; Bergman et al. 1997; Bothe et al. 2010a). So far it is not possible to grow *R. gibba* in axenic cultures (Adler et al. 2010), and for this reason an enrichment culture of the contaminating bacteria was tested for nitrogen fixation. But in this culture no

nitrogenase activity was detected (Floener 1982). Again it was speculated that the cyanobacterial origin of the spheroid bodies mediates the capacity of nitrogen fixation and thereby enables the host's autonomy to combined nitrogen sources (Floener and Bothe 1980). The copy number of spheroid bodies per cell might depend on nitrogen availability in the medium. This could be concluded from a study by DeYoe et al. (1992), indicating that depletion of nitrogen from the medium results in an increasing number of spheroid bodies (DeYoe et al. 1992).

Photosynthesis and nitrogen fixation are mutually exclusive. Thus, diazotrophic cyanobacteria unite both physiological processes either by compartmentalisation or temporal separation. As spheroid bodies might face the same challenge, the spheroid bodies were initially investigated for photosynthesis. But neither chlorophyll autofluorescence nor other photosynthetic pigments could be observed in spheroid bodies (Floener 1982; Kies 1992; Prechtl et al. 2004).

The Origin of the Spheroid Bodies

Prechtl et al. (2004) further examined the *R. gibba* symbiosis. Electron microscopic reexamination of the diatom supported the idea that there is no physical connection between spheroid bodies and the host cytoplasm. In addition, it was shown that the symbiontophoric membrane separates host and endosymbiont, while the spheroid bodies themselves are surrounded by two cyanobacterial envelope membranes (Prechtl et al. 2004). The innermost of these membranes is associated with a thin layer of murein (Drum and Pankratz 1965). In addition, Prechtl et al. (2004) confirmed the light-dependency of nitrogen fixation, as reported earlier (Floener and Bothe 1980).

A breakthrough was the isolation of spheroid bodies. By using Percoll™ density gradients, intact spheroid bodies could be separated from other cell compartments and organelles. Diverse media for cultivating cyanobacteria were applied to grow isolated spheroid bodies, but this approach yielded no results (Prechtl et al. 2004; Kneip et al. 2007).

Isolated spheroid bodies were also used to purify endosymbiont-specific DNA for the amplification of phylogenetically informative sequences. One of the amplificates, the spheroid body-specific SSU rDNA, highlighted homology to cyanobacterial SSU rDNA sequences (Prechtl et al. 2004). A phylogenetic survey indicated that the spheroid body sequence is closely related to nitrogen fixing cyanobacteria of the genus *Cyanothece* (ATCC 51142 and PCC 8801) and to the cyanobacterial endosymbiont of the diatom *Climacodium frauenfeldianum* (Carpenter and Janson 2000; Prechtl et al. 2004; Kneip et al. 2007). Therefore, two conclusions can be made: (a) the spheroid bodies are of cyanobacterial origin and (b) a new model system should be available consisting of related free-living species and intracellular symbionts. Thus, by comparing molecular data, it might be possible to study molecular adaptations necessary for changing from a free-living to an intracellular lifestyle.

Nakayama et al. (2011) demonstrated that spheroid bodies of the Rhopalodiaceae were acquired by a common ancestor. They analysed 18S and 16S ribosomal DNA sequences of *Epithemia turgida*, *Epithemia sorex*, and *R. gibba*. Isolates of all three species were collected at different sites in Japan. Comparison of the 18S rDNA sequences revealed a monophyletic origin of the family Rhopalodiaceae. This clade formed two subclades, one including the *Epithemia* sequences and the other composed of sequences of *R. gibba*. In the 16S rDNA tree of the spheroid bodies, all sequences were recovered as a monophyletic clade. According to the observation of the 18S rDNA tree, this clade was divided into two subclades. One of these subclades corresponded to the *Epithemia* subgroup of the 18S tree, while the second one clearly fitted to the clade of sequences of *R. gibba*. All 16S rDNA sequences were distant from the plastid sequences and showed high similarities to *Cyanothece* species and thereby confirmed the origin of the spheroid bodies (Nakayama et al. 2011).

Localisation of the Nitrogen Fixing Activity

As outlined, the Bothe lab as well as ours determined nitrogen fixing activity of the diatom cell. Thus, it might be likely that this activity is located in the spheroid body. By use of the spheroid body DNA, *nif*D, encoding one of the main components of the nitrogenase enzyme complex, could be identified (Prechtl et al. 2004). NifD is part of the dinitrogenase heterotetramer and is known to be co-transcribed with the other two genes of the nitrogenase in the *nif*HDK operon. Comparison of the *nif*D sequence of the spheroid bodies to those of other diazotrophic organisms again revealed the cyanobacterial ancestry with the highest similarity to the sequence of *Cyanothece* sp. ATCC 51142 (Prechtl et al. 2004; Kneip et al. 2008). To further confirm on the protein level that spheroid bodies are the cellular source of the nitrogen fixation machinery, an antibody against NifD (the α-subunit of the MoFe protein) was used in immuno-gold studies to localise the nitrogenase within the cell. This in situ localisation demonstrated that the nitrogenase is expressed by the spheroid bodies (Prechtl et al. 2004).

Initial Characterisation of the Spheroid Body Genome

Converting lifestyle from a free-living to an intracellular, symbiotic status might lead to drastic physiological reorganisation, which should be programmed in the respective genomes (Fig. 3). Especially for insect–bacteria endosymbioses, it was shown that endosymbiotic genomes are marked by degenerative reorganisation including gene losses, reductions of the genome size in most cases and decreasing G/C contents (Moran et al. 2008).

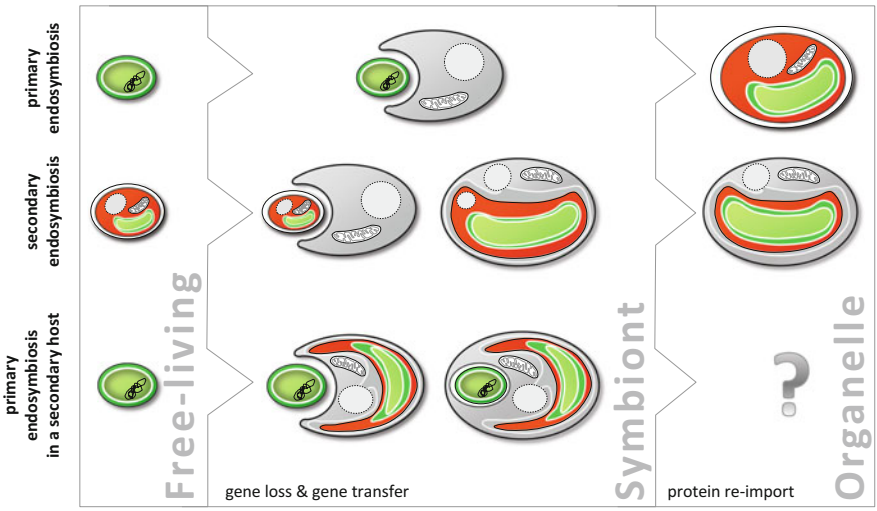


Fig. 3 Schematic illustration of different types of endosymbioses. The engulfment of a cyanobacterium by an eukaryotic host in a primary endosymbiosis finally led to development of primary plastids in glaucophytes, red algae (shown in the *top row*) and green plants. Another recent primary endosymbiosis can be studied in *Paulinella chromatophora*. Secondary endosymbiosis is characterised by the inclusion of an already plastid containing symbiont. The assimilation of a red algae by an eukaryotic host led to the establishment of complex plastids of the heterokontophytes (*middle row*). In the case of *Rhopalodia gibba*, a secondarily photosynthetic eukaryotic host assimilated another cyanobacterium in a (recent) additional primary endosymbiotic event (*last row*). In all different types of endosymbioses the symbiont is characterised by gene losses and/or endosymbiotic gene transfer. The critical step in the transformation from symbiont to an organelle is the establishment of a machinery to reimport the protein products of the respective genes

The physiological capacities of the spheroid bodies described above might indicate either an obligate endosymbiont or a new, nitrogen fixing organelle, which we would like to term in that case "nitrosome." In order to determine the status of the spheroid body, a genome sequencing project was started. Fortunately, the genome sequence of a free-living, near relative *Cyanothece* sp. ATCC 51142 is available, thus allowing a comparative genome project. Such an effort should clarify manifested genome modifications and thereby indicate the status of the spheroid bodies (Fig. 3).

As reported (Kneip et al. 2008), the spheroid body genome project was started by constructing a DNA fosmid library and shotgun sequencing. Thus, by analysing the data already available, first insights in the genome evolution of spheroid bodies can be gained:

Genome Size

The genome size of obligate endosymbionts is generally reduced in comparison to free-living relatives, caused either gradually by gene deletions one by one or by deleting large regions according to reorganisation of chromosomes

(Cole et al. 2001; Moran and Mira 2001; Silva et al. 2001; Toh et al. 2006). In any case, genome reduction might be a time-dependent process. For example, in the case of the ca. 100 My old symbiosis of *Buchnera aphidicola* in aphids (Moran et al. 2008), the size of the symbiont's genome (641 kb) indicates a reduction of 87 % compared to its close relative *Escherichia coli* (Shigenobu et al. 2000). The size of the spheroid body's genome was expected to be smaller than the genome of *Cyanothece* via the loss of genes (Kneip et al. 2008) and estimated to be approximately 2.6 Mb that would imply a reduction of 50 % compared to *Cyanothece* (Kneip et al. 2008; Welsh et al. 2008). Thus, if the estimation of the genome size of the spheroid body is correct, the *R. gibba* spheroid body association can be assumed as a relatively young symbiosis. This is in agreement with the fossil record, indicating that Rhopalodiaceae originated in the Miocene (Hajos 1973; Simonsen 1979; Nakayama et al. 2011).

Pseudogenes

The currently still incomplete sequence information on the spheroid body genome (to be published) indicated that approximately 60 % of the genome is occupied by ORFs; most of them were identified as orthologous sequences of proteins with known or proposed cellular functions. 31 % are conserved hypothetical ORFs and 10 % exhibit no homology to assigned proteins. About 40 % of the spheroid body's genome seems to be noncoding sequences, and in these regions several dozen pseudogenes were identified by internal stop codons or frame shifts.

Pseudogenisation might be useful to calibrate the origin of intracellular symbioses. During streamlining of the symbiontic genomes by losses of genes with functions no longer needed, genes were first inactivated by mutations according to decreasing selection pressure. If the decreased selection pressure leads to a widespread inactivation of genes that depends on each other, this effect is termed "domino theory" (Dagan et al. 2006), thereby first increasing the amount of pseudogenes. However, during further intracellular adaption of the genomes, pseudogenes might become preferred targets for deletions. This might be the reason why in recently established endosymbionts, a large number of pseudogenes are present. In the case of *Sodalis glossinidius*, a symbiont of the tsetse fly *Glossina* spp., 972 pseudogenes were identified (Toh et al. 2006). Contrarily, in comparatively old endosymbiotic interactions as in *B. aphidicola*, only few pseudogenes are detectable (Shigenobu et al. 2000). The already analysed sequences of the spheroid body's genome exhibit several dozens of pseudogenes which again indicate a recently established symbiosis.

Mobile Elements

It is known that symbiotic bacteria, which recently became host-dependent, have a high load of IS elements (Moran and Plague 2004). In contrast, ancient symbioses—with the exception of some strains such as *Wolbachia pipientis* wMel (Moran and Plague 2004; Wu et al. 2004)—have eliminated most of the mobile elements. A clear analysis of IS elements in the spheroid body genome is not possible at the moment. However, many of the identified pseudogenes of the spheroid body's genome are inactivated transposases, which might indicate at least a relatively young spheroid body-diatom symbiosis.

G/C Content

Another attribute of intracellular symbionts is the G/C content of the respective genomes. Genomes of recently established symbionts, such as *S. glossinidius* (Toh et al. 2006), show no pronounced change in G/C content in comparison to free-living relatives. On the other hand, longer lasting symbioses as found in the *Carsonella rudii*/Psyllids association or in the case of *B. aphidicola* show a strong A/T bias compared to free-living relatives (Shigenobu et al. 2000; Nakabachi et al. 2006) with a dominance in noncoding regions and at the third codon position (Bentley and Parkhill 2004). In case of the spheroid body genome, the free-living relatives are already A/T rich. However, the overall G/C-content of the spheroid body genome is, according to the available data, decreased by 3 % in comparison to *Cyanothece* sp. ATCC 51142. Most of the G/C decrease of the spheroid body genome was found in noncoding regions and at the third codon position. A summary of these observations might indicate first steps of a widespread reorganisation of the spheroid body's genome, although a possible A/T-bias might not be a key argument.

DNA Repair

Although not manifested for every intracellular symbiosis (Moran et al. 2008), loss and/or inactivation of genes encoding DNA repair enzymes are further hints of an intracellular lifestyle. Especially in old symbiotic interactions, a tendency for losing the activity of *rec*A and *rec*F can be observed (Moran and Mira 2001; Tamas et al. 2002). However, in the case of the spheroid bodies, the genes for both enzymes are still present.

Thus, the modifications observed in the spheroid body's genome do not finally determine the status of the *R. gibba*/spheroid body association. According to recent

discussions on the definition of organelles and endosymbionts, transfer of genes from the symbiont into the host genome together with the establishment of a protein import machinery to supply the symbiont with the respective gene products might be the clue for a correct decision (Cavalier-Smith and Lee 1985; Keeling and Archibald 2008; Theissen and Martin 2006). If so, the spheroid body might not be an organelle as no transferred genes were detected in a normalised cDNA library of *R. gibba* (unpublished).

The UCYN-A/Spheroid Body Relations

Biological nitrogen fixation has a big evolutionary impact and significantly contributes to the nitrogen cycle in marine habitats [reviewed in Fiore et al. (2010)]. Filamentous cyanobacteria like *Trichodesmium* or the symbiotic *Richelia* as well as unicellular cyanobacteria like *Crocosphaera watsonii* and *Cyanothece* sp. (belonging to the unicellular cyanobacteria group B) are important nitrogen fixing bacteria (Capone et al. 2005; Goebel et al. 2007; Ward et al. 2007; Zehr et al. 2008; Bothe et al. 2010a, b). Interestingly, recent phylogenetic studies on the nitrogenase as well as on 16S rDNA gene sequences (*rrn*16S) of marine communities identified a novel group of cyanobacteria that is closely related to *Cyanothece* sp. ATCC 51142 and the spheroid bodies of *R. gibba* (Zehr et al. 2008; Bothe et al. 2010b). These cyanobacteria, the UCYN-A group (unicellular cyanobacteria group A), are characterised by a small cell size compared to group B (diameter >1 μm) and so far it was not possible to cultivate these organisms (Zehr et al. 2008). Physiological together with molecular investigations indicated that the UCYN-A cyanobacteria lack (partially or completely) several crucial metabolic pathways, including the tricarboxylic acid cycle, the Calvin cycle, biosynthesis of several amino acids and de novo purine biosynthesis (Tripp et al. 2010; Zehr 2011). In addition, this cyanobacterial group has eliminated photosytem II (Zehr et al. 2008; Tripp et al. 2010), a situation similar to the spheroid bodies of *R. gibba*, for which photosystem II pseudogenes were identified (Kneip et al. 2008). The similarity between UCYN-A cyanobacteria and the spheroid bodies is furthermore reflected by the analogous gene arrangement and composition of the nitrogenase gene cluster (Zehr et al. 2008). Moreover, the activity of the nitrogenase in the spheroid bodies and UCYN-A cyanobacteria is light dependent (Floener and Bothe 1980; Prechtl et al. 2004; Zehr et al. 2008). Therefore, adaptations which are thought to be specific for intracellular lifestyle can be recapitulated in free-living organisms.

From Free-Living Cyanobacteria to Endosymbionts to *Paulinella chromatophora* to Plastids?

It seems now possible to reconstruct different steps in the transition from a free-living cyanobacterium to an intracellular endosymbiont, which later might develop into an organelle (Fig. 3). For such a transition line, another primary acquisition of a cyanobacterium can be followed in the cercozoan amoeba *P. chromatophora*. This model is unusual, especially due to the presence of photosynthetically active "chromatophores" (Lauterborn 1895). Initial molecular work on the chromatophores clearly indicated an endosymbiont of cyanobacterial origin, branching with *Synechococcus* WH5701 (Marin et al. 2005). Fortunately, the genomes from two *P. chromatophora* strains were sequenced and analysed (Nowack et al. 2008; Yoon et al. 2009, 2013; Reyes-Prieto et al. 2010), indicating characteristics of intracellular lifestyle such as genome reduction in comparison to the free-living relatives and reduced G/C contents. Moreover, several genes involved in photosynthesis not encoded by the chromatophore were detected in the nuclear genome (Nakayama and Ishida 2009; Reyes-Prieto et al. 2010; Nowack et al. 2011). There are hints about the import of these nuclear-encoded proteins (Bodył et al. 2010; Yoon et al. 2013), but in any case the chromatophore should have the status of an organelle. The *Paulinella*/chromatophore symbiosis is more advanced than the *Rhopalodia*/spheroid body association not only in its molecular characteristics but also in its morphology as shown by the presence of two membranes surrounding the chromatophore instead of three in the diatom example. Thus, independent of the major contribution of the symbiont to the host cell (nitrogen fixation/photosynthesis), a valuable transition line is now available, ranging from free-living cyanobacteria to spheroid bodies to chromatophores to organelles such as primary plastids.

Acknowledgments This work was supported by grants of the Deutsche Forschungsgemeinschaft (DFG SFB TR1). We would like to thank Helen Rennie for critical notes on the manuscript and Magnus Rath for microscopic support.

References

Adler S, Dersch LM, Maier U-G (2010) Cultivation of the diatom *Rhopalodia gibba*. J Endocytobiosis Cell Res 20:109–112

Bentley SD, Parkhill J (2004) Comparative genomic structure of prokaryotes. Annu Rev Genet 38:771–791

Bergman B, Gallon J, Rai A, Stal L (1997) N2 Fixation by non-heterocystous cyanobacteria1. FEMS Microbiol Rev 19:139–185

Bodył A, Mackiewicz P, Stiller JW (2010) Comparative genomic studies suggest that the cyanobacterial endosymbionts of the amoeba *Paulinella chromatophora* possess an import apparatus for nuclear-encoded proteins. Plant Biol 12:639–649

Bolte K, Bullmann L, Hempel F, Bozarth A, Zauner S, Maier U-G (2009) Protein targeting into secondary plastids. J Eukaryot Microbiol 56:9–15

Bothe H, Schmitz O, Yates MG, Newton WE (2010a) Nitrogen fixation and hydrogen metabolism in cyanobacteria. Microbiol Mol Biol Rev 74:529–551

Bothe H, Tripp HJ, Zehr JP (2010b) Unicellular cyanobacteria with a new mode of life: the lack of photosynthetic oxygen evolution allows nitrogen fixation to proceed. Arch Microbiol 192:783–790

Capone DG, Burns JA, Montoya JP, Subramaniam A, Mahaffey C, Gunderson T, Michaels AF, Carpenter EJ (2005) Nitrogen fixation by *Trichodesmium* spp.: an important source of new nitrogen to the tropical and subtropical North Atlantic Ocean. Global Biogeochem Cy 19: GB2024

Carpenter EJ, Janson S (2000) Intracellular cyanobacterial symbionts in the marine diatom *Climacodium frauenfeldianum* (Bacillariophyceae). J Phycol 36:540–544

Cavalier-Smith T, Lee JJ (1985) Protozoa as hosts for endosymbioses and the conversion of symbionts into organelles. J Protozool 32:376–379

Cole S, Eiglmeier K, Parkhill J, James K, Thomson N, Wheeler P et al (2001) Massive gene decay in the leprosy bacillus. Nature 409:1007–1011

Dagan T, Blekhman R, Graur D (2006) The "domino theory" of gene death: gradual and mass gene extinction events in three lineages of obligate symbiotic bacterial pathogens. Mol Biol Evol 23:310–316

DeYoe HR, Lowe RL, Marks JC (1992) Effects of nitrogen and phosphorus on the endosymbiont load of *Rhopalodia gibba* and *Epithemia turgida* (Bacillariophyceae). J Phycol 28:773–777

Drum RW, Pankratz S (1965) Fine structure of an unusual cytoplasmic inclusion in the diatom genus, *Rhopalodia*. Protoplasma 60:141–149

Falkowski PG, Katz ME, Knoll AH, Quigg A, Raven JA, Schofield O, Taylor FJR (2004) The evolution of modern eukaryotic phytoplankton. Science 305:354–360

Fiore CL, Jarett JK, Olson ND, Lesser MP (2010) Nitrogen fixation and nitrogen transformations in marine symbioses. Trends Microbiol 18:455–463

Floener L (1982) Physiologische und biochemische Untersuchungen an dem Cyanellen enthaltenden Flagellat *Cyanophora paradoxa* und an *Rhopalodia gibba*, einer Diatomee mit blau-grünen Einschlüssen. Thesis, University of Cologne

Floener L, Bothe H (1980) Nitrogen fixation in *Rhopalodia gibba*, a diatom containing bluegreenish inclusions symbiotically. In: Schwemmler W, Schenk H (eds) Endocytobiology: endosymbiosis and cell biology, a synthesis of recent research, 1st edn. Walter de Gruyter and Co., Berlin, pp 541–552

Geitler L (1977) Zur Entwicklungsgeschichte der Epithemiaceen: *Epithemia*, *Rhopalodia* und *Denticula* (Diatomophyceae) und ihre vermutlich symbiotischen Sphäroidkörper. Plant Syst Evol 128:259–275

Goebel NL, Edwards CA, Church MJ, Zehr JP (2007) Modeled contributions of three types of diazotrophs to nitrogen fixation at Station ALOHA. ISME J 1:606–619

Hajos M (1973) The Mediterranean diatoms. Initial Rep Deep Sea 13:944–969

Hempel F, Bozarth A, Sommer MS, Zauner S, Przyborski JM, Maier U-G (2007) Transport of nuclear-encoded proteins into secondarily evolved plastids. Biol Chem 388:899–906

Hoek CVD, Jahns HM, Mann DG (1993) Algen, 3rd edn. George Thieme Verlag, Stuttgart

Keeling PJ, Archibald JM (2008) Organelle evolution: what's in a name? Curr Biol 18:R345–R347

Kies L (1992) Glaucocystophyceae and other protists harbouring prokaryotic endocytobionts. In: Reiser W (ed) Algae and symbioses. Biopress Limited, Bristol, pp 353–377

Klebahn H (1896) Beiträge zur Kenntnis der Auxosporenbildung. I. *Rhopalodia gibba*. Jahrbücher der wissenschaftlichen Botanik 29:595–654

Kneip C, Lockhart P, Voss C, Maier U-G (2007) Nitrogen fixation in eukaryotes – new models for symbiosis. BMC Evol Biol 7:55

Kneip C, Voss C, Lockhart PJ, Maier U-G (2008) The cyanobacterial endosymbiont of the unicellular algae *Rhopalodia gibba* shows reductive genome evolution. BMC Evol Biol 8:30

Lauterborn R (1895) Protozoenstudien. II. *Paulinella chromatophora* nov. gen. spec., ein beschalter Rhizopode des Süßwassers mit blaugrünen chromatophorenartigen Einschlüssen. Z Wiss Zool 59:537–544

Marin B, Nowack EC, Melkonian M (2005) A plastid in the making: evidence for a second primary endosymbiosis. Protist 156:425–432

Moran NA, Mira A (2001) The process of genome shrinkage in the obligate symbiont *Buchnera aphidicola*. Genome Biol 2:research0054.1–research0054.12

Moran NA, Plague GR (2004) Genomic changes following host restriction in bacteria. Curr Opin Genet Dev 14:627–633

Moran NA, McCutcheon JP, Nakabachi A (2008) Genomics and evolution of heritable bacterial symbionts. Annu Rev Genet 42:165–190

Moustafa A, Beszteri B, Maier U-G, Bowler C, Valentin K, Bhattacharya D (2009) Genomic footprints of a cryptic plastid endosymbiosis in diatoms. Science 324:1724–1726

Nakabachi A, Yamashita A, Toh H, Ishikawa H, Dunbar HE, Moran NA, Hattori M (2006) The 160-kilobase genome of the bacterial endosymbiont *Carsonella*. Science 314:267

Nakayama T, Ishida K-I (2009) Another acquisition of a primary photosynthetic organelle is underway in *Paulinella chromatophora*. Curr Biol 19:R284–R285

Nakayama T, Ikegami Y, Nakayama T, Ishida K-I, Inagaki Y, Inouye I (2011) Spheroid bodies in Rhopalodiacean diatoms were derived from a single endosymbiotic cyanobacterium. J Plant Res 124:93–97

Nowack EC, Melkonian M, Glöckner G (2008) Chromatophore genome sequence of *Paulinella* sheds light on acquisition of photosynthesis by eukaryotes. Curr Biol 18:410–418

Nowack EC, Vogel H, Groth M, Grossman AR, Melkonian M, Glöckner G (2011) Endosymbiotic gene transfer and transcriptional regulation of transferred genes in *Paulinella chromatophora*. Mol Biol Evol 28:407–422

Pfitzer E (1869) Über den Bau und die Zellteilung der Diatomeen. Botanische Zeitung 27:774–776

Prechtl J, Kneip C, Lockhart PJ, Wenderoth K, Maier U-G (2004) Intracellular spheroid bodies of *Rhopalodia gibba* have nitrogen-fixing apparatus of cyanobacterial origin. Mol Biol Evol 21:1477–1481

Reyes-Prieto A, Yoon HS, Moustafa A, Yang EC, Andersen RA, Boo SM, Nakayama T, Ishida K-I, Bhattacharya D (2010) Differential gene retention in plastids of common recent origin. Mol Biol Evol 27:1530–1537

Round FE, Crawford RM, Mann DG (1990) The diatoms – biology and morphology of the genera. Cambridge University Press, Cambridge, p 125

Shigenobu S, Watanabe H, Hattori M, Sakaki Y, Ishikawa H (2000) Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. APS. Nature 407:81–86

Silva FJ, Latorre A, Moya A (2001) Genome size reduction through multiple events of gene disintegration in *Buchnera* APS. Trends Genet 17:615–618

Simonsen R (1979) The diatom system: ideas on phylogeny. Bacillaria 2:9–71

Tamas I, Klasson L, Canbäck B, Näslund AK, Eriksson A-S, Wernegreen JJ, Sandström JP, Moran NA, Andersson SGE (2002) 50 million years of genomic stasis in endosymbiotic bacteria. Science 296:2376–2379

Theissen U, Martin W (2006) The difference between organelles and endosymbionts. Curr Biol 16:R1016–R1017

Toh H, Weiss BL, Perkin SA, Yamashita A, Oshima K, Hattori M, Aksoy S (2006) Massive genome erosion and functional adaptations provide insights into the symbiotic lifestyle of *Sodalis glossinidius* in the tsetse host. Genome Res 16:149–156

Tripp HJ, Bench SR, Turk KA, Foster RA, Desany BA, Niazi F, Affourtit JP, Zehr JP (2010) Metabolic streamlining in an open-ocean nitrogen-fixing cyanobacterium. Nature 464:90–94

Ward BB, Capone DG, Zehr JP (2007) What's new in the nitrogen cycle ? Oceanography 20:101–109

Welsh EA, Liberton M, Stöckel J, Loh T, Elvitigala T, Wang C et al (2008) The genome of *Cyanothece* 51142, a unicellular diazotrophic cyanobacterium important in the marine nitrogen cycle. PNAS 105:15094–15099

Wu M, Sun LV, Vamathevan J, Riegler M, Deboy R, Brownlie JC et al (2004) Phylogenomics of the reproductive parasite *Wolbachia pipientis w*Mel: a streamlined genome overrun by mobile genetic elements. PLoS Biol 2:0327–0341

Yoon HS, Nakayama T, Reyes-Prieto A, Andersen RA, Boo SM, Ishida K-I, Bhattacharya D (2009) A single origin of the photosynthetic organelle in different *Paulinella* lineages. BMC Evol Biol 9:98

Yoon HS, Yang EC, Qiu H, Bhattacharya D (2013) Photosynthetic *Paulinella*–recapitulation of primary plastid establishment. In: Löffelhardt W (ed) Endosymbiosis. Springer, Vienna

Zehr JP (2011) Nitrogen fixation by marine cyanobacteria. Trends Microbiol 19:162–173

Zehr JP, Bench SR, Carter BJ, Hewson I, Niazi F, Shi T, Tripp HJ, Affourtit JP (2008) Globally distributed uncultivated oceanic N2-fixing cyanobacteria lack oxygenic photosystem II. Science 322:1110–1112

Part IV
Autotrophy as the Driving Force for Endosymbiosis: Secondary and Tertiary Endosymbioses

Chromera et al.: Novel Photosynthetic Alveolates and Apicomplexan Relatives

Marjorie Linares, Dee Carter, and Sven B. Gould

Abstract The Apicomplexa were for long represented only by non-photosynthetic parasites, despite the vast majority of them housing a plastid surrounded by four membranes. The amount of membranes already pointed towards the secondary evolutionary origin of the organelle, and phylogenetic analysis then showed it to be of rhodophyte origin. The discovery of *Chromera velia* now provides the link that connects the parasitic phylum with its algal past. Other chromerids have since been described and within a few years many different research fields have begun to explore this new branch at the bottom of the apicomplexan phylum. We summarize reports from various disciplines and provide an overview of the current topics.

Contents

Background 184
Morphology and Life Cycle Stages of *Chromera velia* 185
Discovery of Other Chromerids 187
Phylogenomic Position 189
Chromera in the Light of the Chromalveolate Hypothesis 190
Plastid Bio- and Photochemistry 191
A Merge of Many Worlds, But What Else Can *Chromera* Offer? 193
References 195

M. Linares • D. Carter
School of Molecular Bioscience, University of Sydney, NSW 2006, Australia

S.B. Gould (✉)
Institute for Molecular Evolution, Heinrich-Heine-University Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany
e-mail: gould@hhu.de

W. Löffelhardt (ed.), *Endosymbiosis*, DOI 10.1007/978-3-7091-1303-5_10,

© Springer-Verlag Wien 2014

Background

In paleontology, a convincing chain of evidence depends on the presence of a good fossil record, each record making up a link in the chain. Some fossils are more informative than others, such as that of the Archaeopteryx or Lucy (AL 288-1), as they are able to connect what had previously been loose ends of an evolutionary chain. The same is true for modern molecular phylogenomics, where sometimes a reliable classification and positioning within a tree is not possible due to incomplete sampling across the range of taxa within the phylogeny. In such cases, identification of a new organism can resolve some of the previous doubts or controversies. The discovery of an inconspicuous alga, *Chromera velia*, provided such a link (Moore et al. 2008). Its importance is obvious, considering how often a taxonomic identification of an extant new organism is published in a high impact journal such as Nature.

In the case of *Chromera*, the reason behind its significance is connected with the sixth word of the abstract in the original paper: *Plasmodium*. We are all well aware of the direct impact of the causative agent of malaria—many hundreds of thousands die every year according to the WHO—and the indirect impact it has on the economy and overall development of the most affected countries, which are mainly in sub-Saharan Africa. As a consequence, the worldwide research budget to support the fight against *Plasmodium* has been, and continues to be, immense. It seems ironic that a $5 mosquito net currently still seems the best and most affordable protection against a *Plasmodium* infection. The ability of the parasite to fool the immune system, and the increasing resistance to what was once successful medication, demands novel drug targets. It is in fact the discovery of such a promising drug target, the vestigial plastid in the apicomplexan lineage [reviewed in McFadden (2011)], that makes the link between the photoautotrophic *Chromera velia* and the parasite *Plasmodium* so significant. The finding of the apicoplast organelle immediately raised the question of the evolutionary ancestry of this parasitic lineage and provided the phylogenomic field with a wealth of work and disputes, some of which are still actively discussed (Gould et al. 2008b; Bodyl et al. 2009; Keeling 2009; Green 2011). Although predicted, it was not until the discovery of *Chromera* that an organism was identified that could link the phylogenomic branch containing the parasitic Apicomplexa with the autotrophic algal world. The initial systematic characterization of *Chromera*, published in 2008, has led to a more in-depth analyses, ranging from phylogenomics to biochemistry all the way to photobiology. An interdisciplinary understanding of this organism is sought in order to make use of its potential as a model organism to study, for example, the individual fates and differential reduction of an endosymbiotic secondary plastid in alveolate organisms. Initial work pursuing this goal is reviewed and discussed here.

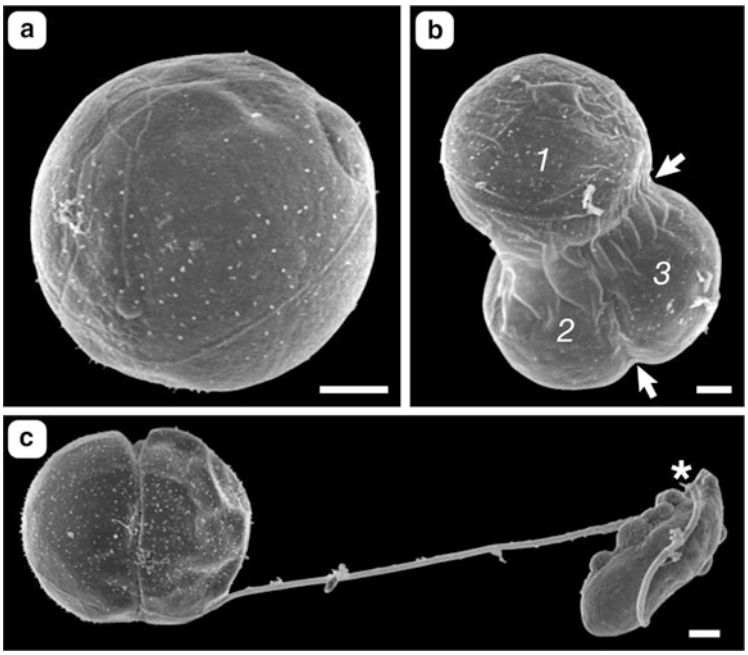


Fig. 1 Scanning electron microscope image of *Chromera velia* stages. (**a**) Single coccoid cell. (**b**) Cyst of coccoid cells containing one single (1) and one double-cell coccoid stage (2 and 3; division furrows marked by an *arrow*). (**c**) Representative of the coccoid stage in bi-cellular form and motile stage (*asterisk*) showing both long and shorter tapered flagella. Scale bar: 1 μm. Pictures courtesy of Kate Weatherby, University of Sydney

Morphology and Life Cycle Stages of *Chromera velia*

Chromera velia was found during a survey for symbiotic dinoflagellates from Pacific stony corals. Initial inspection under the microscope revealed nothing more than an unspectacular, golden-brown unicellular alga (Fig. 1). At face value *Chromera* appeared very similar to the typical and well-studied coral symbiont *Symbiodinium*.

Before the discovery of *Chromera*, the nearest known photosynthetic relatives of Apicomplexa among the higher-ranking superphylum of alveolates were the dinoflagellate algae. Unfortunately, dinoflagellates are always the odd ones out. If you are looking for an exception to the rule—probably any rule—you will find it among dinoflagellates, and in this respect chromerids may offer a much better model for future comparative studies. Upon TEM analysis, features common to alveolates, and in particular to apicomplexan parasites, became apparent. This alga of approximately 5–7 μm in diameter has flattened cortical alveoli with underlying microtubules, a morphological characteristic that was the basis for the first grouping and naming of the Alveolata super group (Cavalier-Smith 1993; Gould et al. 2008a). The presence of a structure called the micropore, an invagination of the plasma

membrane involved in the uptake of nutrients (Nichols et al. 1994; Appleton and Vickerman 1996), also linked *Chromera* with both the dinoflagellate algae and apicomplexans.

Chromera velia has a short-lived flagellated stage that has been characterized by electron microscopy, and this possesses several additional structures characteristic of alveolates (Obornik et al. 2011; Weatherby et al. 2011). The flagellated cells have a smooth cell surface with two heterodynamic flagella, very similar to those of the predatory marine colpodellids. The flagella are fully formed within the cell cytosol and later ejected, a mechanism known as exflagellation and observed in *Plasmodium* (Sinden et al. 2010). A conoid-like structure, also present in perkinsidae and colpodellids—referred to as the pseudoconoid—localizes towards the anterior end, bundling the subpellicular microtubules into a higher ordered structure (Obornik et al. 2011). Analysis of the secondary plastid, tracing back to an engulfed red alga, revealed it to be bound by four membranes, which is characteristic of the apicoplast and different to dinoflagellate plastids, which have three membranes separating the cytosol from the stroma (Gould et al. 2008b; Archibald 2009; Keeling 2010). A unique feature is an apparent organelle that represents an extrusome-like structure, which is not found in apicomplexan parasites and is different from that of dinoflagellates. The function of this feature, dubbed the "chromerosome," remains to be determined, but TEM images showed how it progressed from an early circular form to a striking rod-like protrusion that extends almost across the entire cell (Obornik et al. 2011).

In addition to ultrastructural analysis, Obornik and colleagues investigated how light affects exflagellation in *Chromera*. They found the maximum number of flagellated cells to be present 11 days after subculture, with a peak occurring 6 h into the light period of the diurnal light–dark cycle. Increasing the light intensity resulted in an earlier induction of flagellation to 5 days post-subculture. A study on the effects of different salinity and nutrient concentrations found flagellates to increase in number under conditions that were unfavorable but not toxic, suggesting they play a role in transporting *Chromera* to a more suitable environment (Guo et al. 2010). *Symbiodinium* also has a motile form that occurs during the light phase of the diurnal cycle and that is thought to be used for dispersal or infection to initiate endosymbiosis (Fitt and Trench 1983; Freudenthal 1962). The influence of nutrients, salinity, and light on the motile stage of *Chromera* may indicate that it functions to move cells towards more suitable environments, and perhaps towards new host infection. The appearance of this motile stage at specific points during the diel cycle and upon different intensities of light may point to photosynthetic capacities during the day that can benefit its potential coral host; however, this requires further proof. We do not yet understand the nature of *Chromera*-coral relationship. In this respect it will be of special importance to resolve whether *Chromera* can live endosymbiotically like *Symbiodinium*, or whether it is just loosely associated and free-living, perhaps accessing nutrients that are secreted by the coral or associated biota. We can currently only speculate on the life cycle of *Chromera*, and its true ecological niche and role in the environment have yet to be elucidated.

Intriguingly, some features that are ideal for the parasitic apicomplexan lifestyle seem already hardwired into that of photosynthetic chromerids. As well as the switch from a coccoid to a flagellated form described above, *Chromera* and apicomplexan parasites apparently share a unique form of asexual cell division [reviewed in Striepen et al. (2007)]. Whether it is, for example, endodyogeny (*Toxoplasma*) or schizogony (*Plasmodium*), the aim to rapidly build an efficient, nondividing, and invasive form known as the zoite, involves an intricate daughter cell synthesis process, which begins within the mother cell. Reports on the morphology and ultrastructure of *Chromera*, and its recently described relative *Vitrella brassicaformis* (Obornik et al. 2011, 2012) reveal that in both cases they divide in some kind of "sac," a cyst wall, that encloses dividing cells before their apparent release (Fig. 1). The forming daughter cells of *Toxoplasma gondii* can be visualized with an antibody recognizing alveolins, a family of structural proteins of the subpellicular network, which are associated with the peripheral alveoli (Gould et al. 2008a; Anderson-White et al. 2010). We used this antibody for western blot analysis on total cell lysates of *Chromera*, detecting a single band of approximately 55 kDa (Fig. 2a), and fixed *Chromera* cells to localize the protein through immunofluorescent labeling. Indeed, the localization is reminiscent of that observed in the parasites (Fig. 2b, c). Interestingly, alveolate genomes generally seem to encode only a few copies of this protein family (approximately 1–3 per genome), but apicomplexan parasites encode many more: for example *T. gondii* 14 and *P. falciparum* and the cattle parasite *Theileria parva* each five versions. Should the *Chromera* genome possess only one copy, as revealed by the western blot but has yet to be verified by the genome sequence, this family of cytoskeletal proteins may have expanded in apicomplexan parasites after their separation from the chromerid lineage. However, all of these features seem ancient and to have developed in the common ancestor of Apicomplexa and the chromerid lineage. It is tempting to conclude that such a cloak was a perfect prerequisite for the transition from a mutualistic symbiont to an invading and host-dependent parasite.

Discovery of Other Chromerids

The discovery of *Chromera* has prompted the search for biodiversity within this fascinating group. In the same survey that yielded *Chromera*, another photosynthetic alveolate appeared. Its plastid genome was sequenced and analyzed along with that of *Chromera* (Janouskovec et al. 2010) and will be discussed later in this chapter. The new alveolate has recently been formally described and named *Vitrella brassicaformis* (Oborník et al. 2012). Molecular and ultrastructural analysis of *Vitrella* has found it to be related to apicomplexans, albeit in an apparently independent lineage from *Chromera*. *Vitrella* has a thick laminated, but transparent cell wall, which the authors speculate to aid photosynthesis at ocean depths or even within a coral host. It is further hypothesized to help in surviving harsh environmental changes as well as predation (Oborník et al. 2012). But there are profound

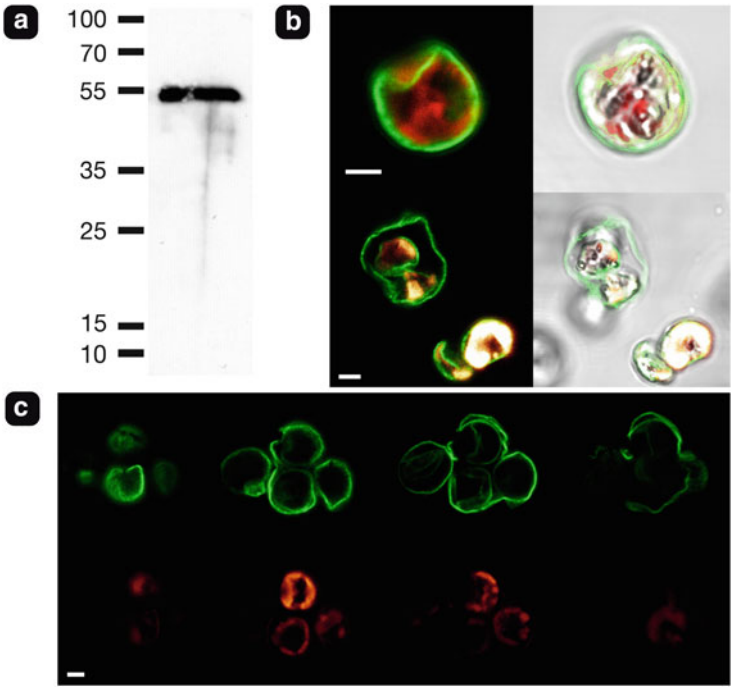


Fig. 2 Alveolin localization in *Chromera velia*. (**a**) Western blot analysis shows *C. velia* to express at least one alveolin protein of approximately 55 kDa. (**b**) The alveolin protein (*green*) localizes to the periphery of the cells and surrounds the plastid auto-fluorescence (*red glow*). On the *right* a merge with the bright field image. (**c**) Series of images showing the association of the protein with the outer layer (possibly the cyst wall) surrounding four individual coccoids. Localization of alveolin in the lower panel of (**b**) and (**c**) is reminiscent of the localization seen to the periphery of the mother cell and newly forming daughter cells in a dividing *Toxoplasma* tachyzoite. Scale bar: 2 μm

differences between *Vitrella* and *Chromera*, in both morphological and genetic terms. One intriguing feature that separates them is that *Vitrella* possesses two distinctly different cell types in different life cycle stages. The first type sporulates to produce nonmotile cells, virtually identical to the vegetative coccoid stage. The second type produces motile bi-flagellated and colorless zoosporangia. These zoosporangia multiply and aggregate within a vacuole and later exit, presumably to move to a new host or to a new environment. *Vitrella* can reach up to 40 μm in diameter, (over 4× that of *Chromera*), and a single sporangium can contain a dozen motile flagellated zoospores, as opposed to one flagellate arising from one coccoid form in *Chromera*. Additionally, *Vitrella* lacks a chromerosome and the distinct finger-like projection from one of its flagella that is seen in *Chromera* (Weatherby et al. 2011; Oborník et al. 2012).

But even with these differences, the similarities between *Chromera* and *Vitrella* are also quite striking and further support their relationship and phylogenomic

position as a lineage basal to apicomplexans (see below). Aside from the typical alveolate features, such as alveoli and undulating thylakoid stacks, the flagellated forms of both organisms have a similar pear shape and heterodynamic flagella, with the flagella tapering towards their terminal end (this is also found in dinoflagellates and colpodellids). In *Vitrella*, mature flagella also form within the cytosol and are later ejected, similar to what is seen for *Chromera* and *Plasmodium* (Sinden et al. 2010). We have only just begun to sample and characterize this novel group of chromerid algae, and the diversity already observed between *Chromera* and *Vitrella* indicates that there is probably much more biodiversity to be discovered at the phylogenomic base of the Apicomplexa.

Phylogenomic Position

The speed with which novel genetic information is being acquired leads to the resolution of some debates but at the same time adds to the complexity of others. Chromerids here are no exception. Initial molecular phylogenetic analyses of just a few canonical marker genes (e.g., LSU and SSU rDNA) showed that *Chromera* is more closely related to apicomplexan parasites than to photosynthetic dinoflagellates, with the new taxon branching at the root of colpodellids (Moore et al. 2008). A multi-gene approach teasing apart all genes in the tetrapyrrole biosynthesis pathway for heme and chlorophyll production not only produced similar evidence that supported the link between *Chromera* and apicomplexans but also suggested that this pathway is a mosaic of genes of various origins (Koreny et al. 2011). Although members of the pathway are encoded in the nucleus, they do not generally display a phylogeny corresponding to that of the host cell, signifying their origin from a plastid genome with later horizontal transfer. *Chromera* does however possess genes for heme biosynthesis that appear to share an origin with those present in the Apicomplexa, giving strong support for a single shared ancestor.

Although an independent lineage is sometimes hard to strictly define—as this definition is a combination of many individual facts and methods—the evidence presented by the phylogenetic analysis of some individual pathways, the concatenated analysis of thousands of nuclear encoded *Chromera* genes (see below), and the identification and analysis of *Vitrella*, it is fair to say that chromerids represent an independent lineage situated between the Perkinsidae and Apicomplexa.

Chromera in the Light of the Chromalveolate Hypothesis

The chromalveolate hypothesis states that all algae that synthesize chlorophyll *c*—that is heterokonts, haptophytes, cryptophytes, and dinoflagellates—as well as the ciliates, apicomplexans, oomycetes, and Perkinsidae, arose from a monophyletic event, i.e., from a single common ancestor that had acquired a plastid of red alga origin by secondary endosymbiosis (Cavalier-Smith 1999). For many years the majority of data appeared to support this hypothesis, but with better sampling and constantly improving phylogenomic methods, this scenario is being challenged. A debate on this topic is well beyond the purpose of this chapter, and we shall therefore focus on the part of chromerids within this enigma.

Sequencing of the two chromerid plastid genomes found them to be smaller than those of heterokonts but significantly larger than those of peridinin-containing dinoflagellates or Apicomplexa. Multi-gene trees derived from the plastid-encoded genes revealed these to be undoubtedly of red algal origin and to position at the base of the apicomplexan branch (Janouskovec et al. 2010). Most red algal-derived plastid genomes contain several conserved gene clusters, including the well-known ribosomal protein superoperon and the atpA-operon, which are also found in apicomplexans, *Vitrella*, and in an interrupted form in *Chromera*. A comparison with a dinoflagellate plastid derived by tertiary endosymbiosis further supported the red algal origin of the *Chromera* plastid (Gabrielsen et al. 2011). However, gene order in the chromerid plastid genome was less conserved than in the *Vitrella* plastid when compared to the apicoplast genome. Strong support for the close relationship between the chromerid, dinoflagellate, and apicomplexan plastids comes from two unique features: (1) replacement of the usually plastid-encoded Rubisco type I by a nuclear-encoded copy of bacterial origin (type II) through horizontal gene transfer (Janouskovec et al. 2010) and (2) unusual codon usage within the *psbA* gene of *Chromera*, where UGA replaces UGG to code for tryptophan at seven conserved positions, which is characteristic of coccidian apicoplasts (Lang-Unnasch and Aiello 1999). Intriguingly, *Vitrella* uses the universal genetic code. Thus, phylogenomic analysis of the plastid genomes gave very strong support to the theory that these are of red algae origin and closely related to those of apicomplexans, dinoflagellates, and heterokonts.

The plastid genome, however, can only tell us something from the endosymbiont perspective and falls short in providing information regarding the eukaryotic host cell. But in the case of the chromalveolate hypothesis, knowing the origin and number of hosts involved is just as crucial. Do the thousands of nuclear-encoded genes that appear to have been transferred during endosymbiont establishment (Martin 2003) point to a single plastid acquisition involving a red alga? A screening of diatom genomes (whose plastid is related to that of alveolates—see above) found an unexpected high amount of genes whose phylogenomic analysis suggest they are of green algal origin (Moustafa et al. 2009) and if so—and if the host is related with that of alveolate organisms—then one must predict to find a similar amount of "green genes" among chromerids. Woehle et al. (2011) performed a concatenated

analysis of more than 3,000 ESTs from *Chromera* and found 513 nuclear-encoded genes with a homolog in red as well as green algae. Of these, 263 appeared to be of green algal origin. The authors provided two scenarios, the simplest, to explain their finding. The first includes an ancestral independent establishment of secondary endosymbiosis with a green algal cell, followed by elimination of the green plastid through the acquisition of a red algal endosymbiont. The second scenario comprises a green algal endosymbiosis that had previously been established by a red algal ancestor, before this red algae became an endosymbiont itself: within both chromerids and diatoms. Along with proposing these scenarios, the authors searched for an explanation easier to "digest" and tested for possible phylogenetic error. They found that in this particular case sampling bias, that is a lack of genomic information for red algae compared to green algae, could cause their observation. This study illustrated how careful one must be when interpreting these kinds of data, and proposing evolutionary scenarios to explain extant species diversity, even if the latest phylogenomic methods are used.

Plastid Bio- and Photochemistry

The link between the apicoplast and the *Chromera* plastid, as well as shedding light on the evolution and loss of photosynthesis, may also reveal novel drug targets within conserved biochemical pathways. To date, heme, chlorophyll, and lipid biochemistry of the *Chromera* plastid have been studied and unique aspects of *Chromera* photochemistry been revealed.

Heme biosynthesis, which uses the same pathway as the production of chlorophyll, was one of the first pathways studied to clarify the apicomplexan-chromerid link. Heme synthesis in *Chromera* operates in a similar fashion to that in animals and fungi (and apicomplexans), via the C4-pathway starting from glycine and succinyl-CoA, as opposed to glutamate, the commonly used starting point for phototrophs. Koreny et al. (2011) cultured *Chromera* in the presence of both radiolabeled glycine (^{14}C-Gly) and glutamate (^{14}C-Glu), and could track the incorporation of only ^{14}C-Gly into chlorophyll, indicating a parallel to that of parasites. Yet in contrast to apicomplexans, *Chromera* uses four different enzymes in each step of heme production, which are found in other photosynthetic eukaryotes. Additionally, and perhaps most interestingly, the ferrochelatase catalyzing the last step in heme synthesis appears to have been acquired from a proteobacterial ancestor of *Chromera* and apicomplexans. Moreover, an apparent compartmentalization of components of the heme and chlorophyll production pathway, where precursors are located in the mitochondrion and their products are translocated to the plastid where chlorophyll and heme are formed, appears unique to *Chromera*. In the case of apicomplexan parasites, this process is cyclical: precursors (glycine and succinyl CoA) form early pathway intermediates in the mitochondrion; these are then translocated to the plastid for the subsequent synthesis of later intermediates (which in some cases also occurs partly in the cytosol),

and then returned to the mitochondrion for heme production. As photosynthetic eukaryotes generally produce higher levels of chlorophyll compared to heme, this entire process is carried out in the plastid, whereas in heterotrophic eukaryotes and parasites, the bulk of heme is produced and used by the mitochondrion for energy purposes. Current data suggests *Chromera* has combined several pathways, merging independent sources, to set up its own unique heme-processing pathway.

Galactolipids are the most abundant form of lipids in membranes of plants and algal plastids, making up around 85 % of the lipid mass. Monogalactosyldiacylglycerol (MGDG) and digalactosyldiacylglycerol (DGDG) are produced by two transferases localized in the outer plastid membrane (Joyard et al. 2010; Botté et al. 2011). They are presumed to be tightly associated with the photosystems and light-harvesting complexes. No galactolipid synthesis genes have been found in Apicomplexa, despite the fact that MGDG and DGDG synthesis have been detected in *P. falciparum* and *T. gondii* (Marechal et al. 2002). The biochemical composition of *Chromera* galactolipids was recently analyzed and found to be equivalent to plant galactoglycerolipids, using spinach as a comparative example.

MGDG synthesis in plants is catalyzed by multiple transferases that modify diacylglycerol (DAG) and attach functional groups using galactosyl, whereas cyanobacteria use glucosylation of DAG, which then undergoes a conformational change to form MGDG. *Chromera* does not use the cyanobacterial glucosylation process for MGDG production and operates more similarly to higher plants, but the plant-like MGDG synthase gene seems derived from bacteria other than cyanobacteria and is likely to have been acquired by horizontal gene transfer (Botté et al. 2011). Their phylogenetic analysis of the MGDG and DGDG synthase homologues in *Chromera* support their grouping with other organisms carrying secondary plastids of red algal origin. To summarize this could mean eukaryotes were able to replace the cyanobacterial two-step synthesis of MGDG with a perhaps more efficient single step, requiring a single enzyme while maintaining the essential end product.

We have discussed above the morphological changes of *Chromera* in response to light, as well as some initial identification of its pigment composition, but additional attention has been paid to more specific photo-responsive capabilities of this alga. Non-photochemical quenching (NPQ) is a mechanism that buffers energy peaks generated during high light stress. It reduces the amount of energy reaching photosynthetic reaction centers, protecting against potential photoinhibition. NPQ occurs in light harvesting antennae and is triggered by the reversible interconversion of xanthophyll pigments via de-epoxidation/epoxidation carried out in a two-step process. The photosynthetic pigment composition of *Chromera* consists of only chlorophyll *a*, as well as violaxanthin and a novel isofucoxanthin (Kotabova et al. 2011). The reaction for NPQ, involving the two-step de-epoxidation of violaxanthin, occurs surprisingly fast in *Chromera* with its intermediate virtually absent, indicating NPQ to be significantly faster in *Chromera* compared to higher plants or heterokont algae (Horton et al. 2008; Garcia-Mendoza and Colombo-Pallotta 2007). Consequently, one can conclude

that *Chromera* seems to use violaxanthin considerably more for photo protection via NPQ, as opposed to the more well-known role in light harvesting.

Further insights into *Chromera* photochemistry were obtained by analysis of expressed genes encoding plant light harvesting complexes (LHC). While some of these grouped with known LHCs from other photosynthetic protists, the vast majority formed a unique clade on a strongly supported monophyletic branch that was closest to fucoxanthin chlorophyll *a*/*c*-binding LHCs from chlorophyll *c*-containing algae, including diatoms, brown algae, and dinoflagellates (Pan et al. 2012). Again, this highlights the unique nature of the biochemistry and photobiology of *Chromera* and its ability to reveal hitherto unrecognized diversity in many eukaryotic systems.

From the analysis of the biochemistry/photochemistry alone, we can currently conclude that *C. velia* is rather odd and almost always seems to differ, at least in some detail, from the systems so far analyzed. In this respect this alga opens many new windows onto the diversity of systems in the eukaryotic world. This again highlights how little we actually know and how much remains to be discovered from currently unknown organisms inhabiting unusual ecological niches. *Chromera velia* being such an oddball makes it just as valuable to us for future studies, as it being a potential new model system and alternative resource to study apicomplexan parasites biology.

A Merge of Many Worlds, But What Else Can *Chromera* Offer?

From the studies outlined above, it is apparent that the chromerid lineage unites many features that are found separately in phylogenetically related species. Now focusing on key topics offers the chance to unravel some of the puzzles that are associated with other alveolates. The identification of *Chromera* has solved a long dispute regarding the origin of the apicoplast and has traced it back to a red algal lineage beyond any doubt. At the same time it has opened a wide field of new research across many disciplines.

In 2008 Okamoto and McFadden discussed the discovery of *Chromera* in a comment they titled “The mother of all parasites” (Okamoto and McFadden 2008). One interesting aspect highlighted by the authors was the evolutionary timescale for the origin and rise of Apicomplexa and their hosts. According to current data, both seem to have been evolving in parallel for more than 500 million years. Between some of them symbiosis developed, and at one point the balance shifted towards parasitism for a few. For the apicomplexan lineage, it appears this transition was a major evolutionary success, leading to the diversity of parasites this group represents today. Their long coevolution with the earliest animals also made apicomplexans experts in evading immune attack. Maybe the focus on parasitic Apicomplexa has diverted our perspective in such a way that we have missed

nonparasitic forms, for example, chromerid-like symbionts of other invertebrates. It will be interesting to see to what degree chromerids, many of which must still await discovery, have kept their ability to shift between a free-swimming phototroph and an intracellular symbiont, and maybe even a fully fledged parasite.

Will chromerids live up to their promise as model organisms for studying parasitism, photosynthesis, and comparative cell biology in Apicomplexa? It appears so. It is likely that further work on this alga, or maybe on other close relatives, will place it as a model organism to study unique apicomplexan features and their evolutionary fate during the transition from a free-living phototroph to an obligatory heterotrophic parasite. An important question that has yet to be answered is whether *Chromera* is a true endosymbiont. Although it was isolated from coral tissue samples, it has not yet been visualized inside a vacuole within a coral cell. Confirming an endosymbiotic relationship will greatly enhance the evidence for an ancestral link with apicomplexan parasites. The first steps have been taken in developing tools using fluorescence in situ hybridization (Morin-Adeline et al. 2012) to answer the question of *Chromera*'s niche within corals.

The phylogenomic relationship of *Chromera* with parasites, and the ease with which it can be cultured in the laboratory, remain compelling reasons for using it as a model organism. *Chromera* cultures reach a density of ~2×10^6 cells/ml within about 2 weeks in inexpensive f/2 medium (Guillard and Ryther 1962). Storage is rather simple: we have 6-month-old cultures left at room temperature without any special light conditions that can still be used for inoculation of fresh media. However, we are not aware of anyone successfully re-culturing *Chromera* from thawed cryostocks. Further technical advances that will enhance the use of *Chromera* will be the availability of a sequenced genome and a transfection protocol. Sequencing of the nuclear genome, currently estimated to be at least 150 Mb, has commenced. Some of the many important questions that the genome sequence will allow us to address include (1) what impact the transition to parasitism has on apicomplexan genomes, apart from losing photosynthesis; (2) what protein families and "apicomplexan orphan genes" shape this lineage; and (3) to what degree are dinoflagellate or perkinsid-like features present next to "apicomplexan features." An efficient transfection protocol is required for the detailed analysis of *Chromera* genes and gene expression, and may also open the way for the heterologous expression of parasite genes, which can be very difficult in current host–vector systems. Once a suitable vector has been generated, with *Chromera*-specific promoters and appropriate selectable markers, the thick cell wall of *Chromera* might present the toughest obstacle when attempting to introduce the foreign DNA. The use of a biolistic gun, widely used for the transfection of plant tissue and diatom algae, might offer a solution.

There is a breadth and depth of tools available to study this intriguing alga from all angles and disciplines. The insight from the studies on *Chromera* to date has given us multiple stepping stones to continue solving the alveolate puzzle. Photobiology has provided us with a better understanding of photochemical and physiological processes within the algal cell to understand more about its natural environment. Phylogenomics placed chromerids within the tree of life and

unexpectedly gave us a peek at the evolution of symbiosis and parasitism once we saw its close relationship with apicomplexans. Biochemistry introduced us to the mixed nature of different biosynthetic pathways, some similar to plants while others more similar to parasites. At every turn we are surprised by the biology of chromerids, and these data have just scratched the surface. *Chromera velia* itself might not be the mother of all parasites, but it surely is a close relative with whom you want to keep in touch.

Acknowledgments This work was supported by the SFF fund of the HHU Düsseldorf and an DFG grant to SBG, an ERC grant to William F. Martin and by an Australian Research Council grant to DC (Discovery Project DP0986372). ML is supported by the William Murrell Memorial Scholarship.

References

Anderson-White BR, Ivey FD, Cheng K, Szatanek T, Lorestani A, Beckers CJ, Ferguson DJ, Sahoo N, Gubbels MJ (2010) A family of intermediate filament-like proteins is sequentially assembled into the cytoskeleton of *Toxoplasma gondii*. Cell Microbiol 13:18–31

Appleton PL, Vickerman K (1996) Presence of apicomplexan-type micropores in a parasitic dinoflagellate, *Hematodinium sp*. Parasitol Res 3:279–282

Archibald JM (2009) The puzzle of plastid evolution. Curr Biol 19:R81–R88

Bodyl A, Stiller JW, Mackiwwicz P (2009) Chromalveolate plastids: direct descent or multiple endosymbioses? Trends Ecol Evol 24:119–121

Botté CY, Yamaryo-Botté Y, Janouskovec J, Rupasinghe T, Keeling PJ, Crellin P, Coppel RL, Maréchal E, McConville MJ, McFadden GI (2011) Identification of plant-like galactolipids in *Chromera velia*, a photosynthetic relative of malaria parasites. J Biol Chem 286:29893–29903

Cavalier-Smith T (1993) Kingdom Protozoa and its 18 phyla. Microbiol Rev 57:953–994

Cavalier-Smith T (1999) Principles of protein and lipid targeting in secondary symbiogenesis: euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree. J Eukaryot Microbiol 46:347–366

Fitt WK, Trench RK (1983) Endocytosis of the symbiotic dinoflagellate *Symbiodinium microadriaticum* Freudenthal by endodermal cells of the scyphistomae of *Cassiopeia xamachana* and resistance of the algae to host digestion. J Cell Sci 64:195–212

Freudenthal HD (1962) *Symbiodinium* gen. nov. and *Symbiodinium microadriaticum* sp. nov., a Zooxanthella: taxonomy, life cycles and morphology. J Protozoology 9:45–52

Gabrielsen TM, Minge MA, Espelund M, Tooming-Klunderud A, Patil V, Nederbragt AJ, Otis C, Turmel M, Shalchian-Tabrizi K, Lemieux C, Jakobsen KS (2011) Genome evolution of a tertiary dinoflagellate plastid. PLoS One 6:e19132

Garcia-Mendoza E, Colombo-Pallotta MF (2007) The giant kelp *Macrocystis pyrifera* presents a different nonphotochemical quenching control than higher plants. New Phytol 173:526–536

Gould SB, Tham WH, Cowman AF, McFadden GI, Waller RF (2008a) Alveolins, a new family of cortical proteins that define the protist infrakingdom Alveolata. Mol Biol Evol 25:1219–1230

Gould SB, Waller RF, McFadden GI (2008b) Plastid evolution. Annu Rev Plant Biol 59:491–517

Green BR (2011) After the primary endosymbiosis: an update on the chromalveolate hypothesis and the origins of algae with Chl *c*. Photosynth Res 107:103–115

Guillard RRL, Ryther JH (1962) Studies of marine planktonic diatoms. I. *Cyclotella nana* Hustedt and *Detonula confervacea* Cleve. Can J Microbiol 8:229–239

Guo JT, Weatherby K, Carter D, Slapeta J (2010) Effect of nutrient concentration and salinity on immotile–motile transformation of *Chromera velia*. J Eukaryot Microbiol 57:444–446

Horton P, Johnson MP, Perez-Bueno ML, Kiss AZ, Ruban AV (2008) Photosynthetic acclimation: does the dynamic structure and macro-organisation of photosystem II in higher plant grana membranes regulate light harvesting states? FEBS J 275:1069–1079

Janouskovec J, Horák A, Oborník M, Lukes J, Keeling PJ (2010) A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids. Proc Natl Acad Sci USA 107:10949–10954

Joyard J, Ferro M, Masselon C, Seigneurin-Berny D, Salvi D, Garin J, Rolland N (2010) Chloroplast proteomics highlights the subcellular compartmentation of lipid metabolism. Prog Lipid Res 49:128–158

Keeling PJ (2009) Chromalveolates and the evolution of plastids by secondary endosymbiosis. J Eukaryot Microbiol 56:1–8

Keeling PJ (2010) The endosymbiotic origin, diversification and fate of plastids. Philos Trans R Soc Lond B Biol Sci 365:729–748

Koreny L, Sobotka R, Janouskovec J, Keeling PJ, Obornik M (2011) Tetrapyrrole synthesis of photosynthetic chromerids is likely homologous to the unusual pathway of apicomplexan parasites. Plant Cell 23:3454–3462

Kotabova E, Kan R, Jarešová J, Prášil O (2011) Non-photochemical fluorescence quenching in *Chromera velia* is enabled by fast violaxanthin de-epoxidation. FEBS Lett 585:1941–1945

Lang-Unnasch N, Aiello DP (1999) Sequence evidence for an altered genetic code in the *Neospora caninum* plastid. Int J Parasitol 29:1557–1562

Marechal E, Azzouz N, de Macedo CS, Block MA, Feagin JE, Schwarz RT, Joyard J (2002) Synthesis of chloroplast galactolipids in apicomplexan parasites. Eukaryot Cell 1:653–656

Martin W (2003) Gene transfer from organelles to the nucleus: frequent and in big chunks. Proc Natl Acad Sci U S A 100:8612–8614

McFadden GI (2011) The apicoplast. Protoplasma 248:641–650

Moore RB, Oborník M, Janouskovec J, Chrudimský T, Vancová M, Green DH, Wright SW, Davies NW, Bolch CJ, Heimann K, Slapeta J, Hoegh-Guldberg O, Logsdon JM, Carter DA (2008) A photosynthetic alveolate closely related to apicomplexan parasites. Nature 451:959–963

Morin-Adeline V, Foster C, Slapeta J (2012) Identification of *Chromera velia* by fluorescence in situ hybridization. FEMS Microbiol Lett 328:144–149

Moustafa A, Beszteri B, Maier UG, Bowler C, Valentin K, Bhattacharya D (2009) Genomic footprints of a cryptic plastid endosymbiosis in diatoms. Science 324:1724–1726

Nichols BA, Chiappano ML, Pavesio CEN (1994) Endocytosis at the micropore of *Toxoplasma gondii*. Parasitol Res 80:91–98

Obornik M, Vancová M, Laia DH, Janouskovec J, Keeling PJ, Lukes J (2011) Morphology and ultrastructure of multiple life cycle stages of the photosynthetic relative of apicomplexa, *Chromera velia*. Protist 162:115–130

Oborník M, Modrý D, Lukeš M, Cernotíková-Stříbrná E, Cihlář J, Tesařová M, Kotabová E, Vancová M, Prášil O, Lukeš J (2012) Morphology, ultrastructure and life cycle of *Vitrella brassicaformis* n. sp., n. gen., a novel chromerid from the great barrier reef. Protist 163:306–323

Okamoto N, McFadden GI (2008) The mother of all parasites. Future Microbiol 3:391–395

Pan H, Slapeta J, Carter D, Chen M (2012) Phylogenetic analysis of the light-harvesting system in *Chromera velia*. Photosynth Res 111:19–28

Sinden RE, Talman A, Marques SR, Wass MN, Sternberg MJE (2010) The flagellum in malarial parasites. Curr Opin Microbiol 13:491–500

Striepen B, Jordan CN, Reiff S, van Dooren GG (2007) Building the perfect parasite: cell division in apicomplexa. PLoS Pathog 3:e78

Weatherby K, Murray S, Carter D, Slapeta J (2011) Surface and flagella morphology of the motile form of *Chromera velia* revealed by field-emission scanning electron microscopy. Protist 162:142–153

Woehle C, Dagan T, Martin WF, Gould SB (2011) Red and problematic green phylogenetic signals among thousands of nuclear genes from the photosynthetic and apicomplexa-related *Chromera velia*. Genome Biol Evol 3:1220–1230

Nucleomorph Comparative Genomics

Goro Tanifuji and John M. Archibald

Abstract Nucleomorphs are vestigial nuclei of endosymbiotic origin found in cryptomonad and chlorarachniophyte algae. A wealth of molecular and comparative genomic data has been generated in recent years providing insight on the origins and evolution of these peculiar organelles. The cryptomonad nucleomorph (and its associated plastid) has been shown to be the product of a secondary (i.e., eukaryote–eukaryote) endosymbiotic event involving a red alga and a heterotrophic host, while chlorarachniophytes have a green algal-derived nucleomorph. Despite their independent origins, the nucleomorphs of both lineages show similar features, most notably the presence of three linear chromosomes and sub-telomeric ribosomal DNA operons. Recent study has revealed similarities between the cryptomonad and chlorarachniophyte nucleomorphs not only in their genome structures but also in their coding content. Significant differences nevertheless exist. For example, spliceosomal introns are rare (or completely absent) in cryptomonad nucleomorph genomes but highly abundant in chlorarachniophytes. In this chapter, we review the current state of knowledge of nucleomorph genome biology, focusing on the evolution, diversity, and function of nucleomorphs in the two lineages that bear them.

Contents

Introduction 198
Nucleomorphs—Remnant Nuclei of Endosymbiotic Origin 200
Convergent Evolution and "The Rule of Three" 201
Nucleomorph Genomes are "Jam-Packed" 201
Nucleomorph Gene Content 204

G. Tanifuji • J.M. Archibald (✉)
Canadian Institute for Advanced Research, Integrated Microbial Biodiversity Program, Department of Biochemistry and Molecular Biology, Dalhousie University, Sir Charles Tupper Medical Building, 5850 College Street, Halifax, NS B3H 4R2, Canada
e-mail: john.archibald@dal.ca

W. Löffelhardt (ed.), *Endosymbiosis*, DOI 10.1007/978-3-7091-1303-5_11,

© Springer-Verlag Wien 2014

ORFan Genes in Nucleomorph Genomes 205
Nucleomorphs: End Result or Intermediate State? 207
Conclusions and Future Directions 209
References 210

Introduction

Endosymbiosis is a driving force in the evolution of eukaryotic cells. All known eukaryotes contain mitochondria (or derivatives thereof), which evolved from bacterial endosymbionts related to modern-day α-proteobacteria (Dolezal et al. 2006; Gray et al. 1999), whereas the plastids (chloroplasts) of algae and plants evolved from once free-living cyanobacteria (Gould et al. 2008; Reyes-Prieto et al. 2007). As described by Gagat et al. (2013) in this volume, photosynthetic eukaryotes have acquired their light-harvesting organelles in several different ways. Red, green, and glaucophyte plastids are believed to be the product of a "primary" endosymbiosis between a eukaryote and a cyanobacterium. Both red and green algal "primary" plastids have subsequently spread to other eukaryotes by "secondary" endosymbiosis, i.e., the engulfment of a primary plastid-bearing alga by a non-photosynthetic host. Examples of tertiary endosymbioses have also been documented in which a secondary-plastid-bearing alga is ingested by another eukaryote, which may or may not itself harbor a plastid (Hackett et al. 2004).

An important aspect of the evolutionary transition from endosymbiont to organelle is the phenomenon of endosymbiotic gene transfer (EGT). Plastids and mitochondria retain at most 5 % of the genetic material thought to have been present in their bacterial progenitors (Martin et al. 2002): this massive reduction is the combined result of EGT and the outright loss of genes unnecessary for an endosymbiotic lifestyle (Martin and Herrmann 1998; Martin 2003). Yet despite their limited coding capacity, plastids and mitochondria still require 1,000 or more proteins to maintain function. During the course of evolution, many (but not all) of the genes that have moved from the endosymbiont to the nucleus have acquired suitable promoter sequences allowing their expression as well as sequences capable of targeting their protein products back to the endosymbiont compartment (Bock and Timmis 2008; van Dooren et al. 2001). Comparative genomics has revealed that the "reorganization" of both the host and endosymbiont genomes that accompanies the evolution of an organelle is far more complex than previously imagined. Endosymbiont-derived proteins have acquired functions specifically related to the host and, conversely, host-derived proteins have been shown to function in mitochondria and plastids [e.g., Martin et al. (1996, 2002) and Martin and Schnarrenberger (1997)].

The "nucleomorphs" of cryptomonad and chlorarachniophyte algae represent a long-standing puzzle in the field of endosymbiosis (Archibald 2007; Archibald and Lane 2009; Cavalier-Smith 2002; Gilson and McFadden 2002; Moore and

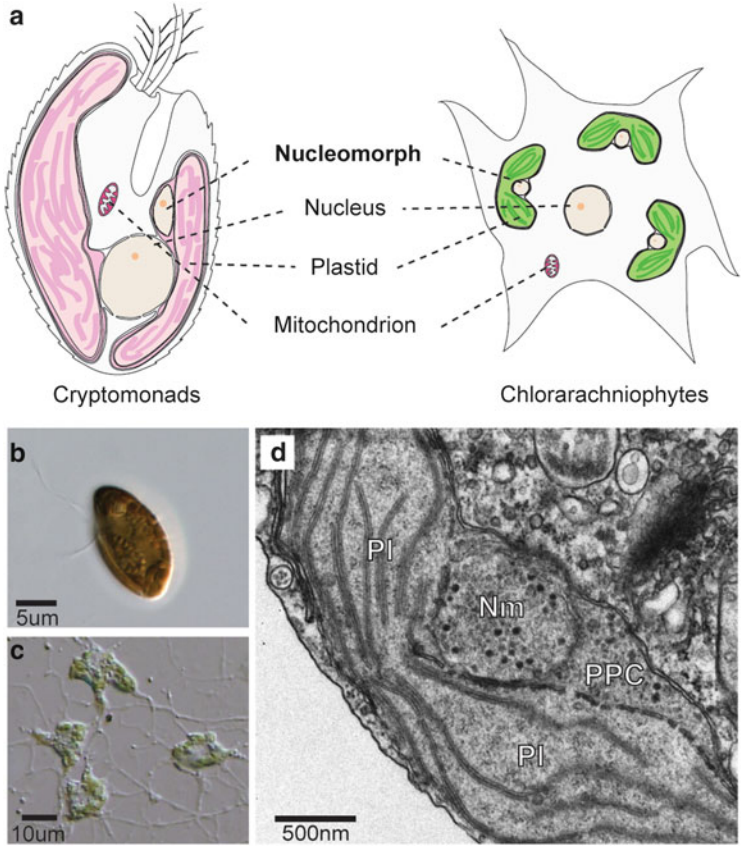


Fig. 1 Basic features of cryptomonads and chlorarachniophytes. (**a**) Schematic diagrams of cryptomonad and chlorarachniophyte cells highlighting the four genome-containing compartments. (**b**) Light micrograph of the cryptomonad *Cryptomonas* sp. (photosynthetic species). (**c**) Light micrograph of the chlorarachniophyte *Chlorarachnion reptans*. (**d**) Transmission electron micrograph of the chlorarachniophyte *Lotharella oceanica*. The plastid (Pl), nucleomorph (Nm), and periplastidial compartment (PPC) are labeled. Chlorarachniophyte images were taken by Dr. Shuei Ota

Archibald 2009) (Fig. 1a, b, c). Nucleomorphs are the residual nuclei of secondary endosymbionts whose persistence is intriguing given that other algae, such as haptophytes and diatoms, clearly have a plastid of secondary endosymbiotic origin but do not retain a nucleomorph. Cryptomonads and chlorarachniophytes are thus interesting organisms in which to study "evolution in action." In this chapter, we review the current state of knowledge of nucleomorph biology focusing on advances coming from comparative genomics. As we shall see, in some ways nucleomorph genomes are extraordinarily dynamic and fast evolving, while in others they appear to be essentially "frozen."

Nucleomorphs—Remnant Nuclei of Endosymbiotic Origin

The nucleomorph of cryptomonad algae was first described by Greenwood (1974). With the use of transmission electron microscopy, researchers at this time had begun to recognize the significant degree to which plastids differed from species to species. Double membrane-bound plastids were found in organisms such as higher plants and the green alga *Chlamydomonas*, while other lineages (e.g., euglenoids, heterokonts, and cryptomonads) had plastids surrounded by three or four membranes [e.g., Gibbs (1978) and Whatley and Whatley (1981)]. Consideration of plastid pigment composition led to speculation that plastids had been acquired by independent endosymbiosis events in multiple lineages [e.g., Gray and Doolittle (1982)]. Plastids surrounded by two membranes are now believed to be the result of an endosymbiosis between a eukaryote and a prokaryote (primary endosymbiosis), while the three or four membrane-bound plastids are the products of symbioses between two eukaryotic cells (secondary endosymbiosis) (Gould et al. 2008; Reyes-Prieto et al. 2007). The cryptomonad nucleomorph was found to exist in the space between the second and third plastid membranes, which corresponds to the cytosol of the eukaryotic endosymbiont (periplastidial compartment; PPC) (Fig. 1d). Nucleomorphs were found to contain DNA (Ludwig and Gibbs 1985) and were also found in another eukaryotic lineage, the chlorarachniophytes (Hibberd and Norris 1984). Together, cryptomonads and chlorarachniophytes are the only two lineages in which these remarkable organelles have been found.

The use of molecular biology in the 1990s revealed that the host nuclear genome and nucleomorph genome encode evolutionarily distinct rDNA sequences, analyses of which revealed that the endosymbiont of cryptomonads was a red alga (Cavalier-Smith et al. 1996; Douglas et al. 1990; Douglas and Penny 1999) and a green alga in chlorarachniophytes (Ishida et al. 1997, 1999; McFadden et al. 1995). Recent advances in our understanding of the eukaryotic tree of life have revealed that the cryptomonad and chlorarachniophyte host lineages belong to two different "supergroups": the "chromalveolates" and Rhizaria, respectively (Burki et al. 2009; Hampl et al. 2009). Thus, the host and endosymbiont components of cryptomonads and chlorarachniophytes are clearly different. Nevertheless, the nucleomorph genomes of both lineages show similar features, raising interesting questions about the evolutionary pressures that have given rise to them. In addition, cryptomonads contain three distinct types of organisms, including photosynthetic and nucleomorph-containing species (e.g., *Guillardia, Rhodomonas, Chroomonas*), non-photosynthetic and nucleomorph-containing species (*Cryptomonas paramecium*), and non-photosynthetic and nucleomorph-lacking species, the latter organisms belonging to a single yet diverse genus (*Goniomonas*) (Deane et al. 2002; Hoef-Emden et al. 2002; Hoef-Emden 2005; McFadden et al. 1994; Shalchian-Tabrizi et al. 2008). The study of nucleomorph genome biology has provided novel insight into the process of reductive evolution and how these different cryptomonads might be related to one another.

Convergent Evolution and "The Rule of Three"

Nucleomorph karyotypic analyses have been carried out for ~50 species of cryptomonads and ~20 species of chlorarachniophytes, using a combination of pulsed-field gel electrophoresis and Southern hybridization (Eschbach et al. 1991; Ishida et al. 2011a; Lane et al. 2006; Lane and Archibald 2006; Phipps et al. 2008; Rensing et al. 1994; Silver et al. 2007; Tanifuji et al. 2010). These data show that all known nucleomorph genomes in both cryptomonads and chlorarachniophytes are comprised of three linear chromosomes ranging from ~100 to ~300 kilobase pairs (kbp). As well, they almost always possess ribosomal DNA operons on the sub-telomeric regions of their chromosomes, although their precise arrangement and content can vary (e.g., in cryptomonads some chromosome termini possess only a 5S rRNA gene) (Lane and Archibald 2006, 2008; Silver et al. 2010). These structural similarities are intriguing given that the nucleomorphs of cryptomonads and chlorarachniophytes do not share common ancestry. Cavalier-Smith suggested that the apparently universal three-chromosome architecture of nucleomorph genomes is related to how these chromosomes are squeezed into the tiny confines of the nucleomorph (Cavalier-Smith 2002). Interestingly, the microsporidian *Encephalitozoon cuniculi*, an obligate intracellular parasite with a genome of only 2.9 megabase pairs (Mbp), also contains sub-telomeric rDNA operons on each of its chromosomes (Katinka et al. 2001). The biological significance of this feature is unknown.

Nucleomorph genome size ranges between 450–865 kbp in cryptomonads and 330–1,000 kbp in chlorarachniophytes, the smallest eukaryotic nuclear genomes known (Ishida et al. 2011a; Moore and Archibald 2009; Silver et al. 2007; Tanifuji et al. 2010) (Fig. 2). Even the smallest genomes of free-living algae, such as members of the Cyanidiophyceae (red algae) and the green algae *Ostreococcus tauri* and *Micromonas* sp., are 10–20 Mbp in size (Derelle et al. 2006; Matsuzaki et al. 2004; Moreira et al. 1994; Muravenko et al. 2001; Worden et al. 2009). Although the extant algae that are most closely related to the progenitors of the cryptomonad and chlorarachniophyte nucleomorphs are not known, these data suggest that the nucleomorph genomes of the two groups have been reduced to <10 % of that of the free-living algae from which they evolved. Comparative genomics has made it possible to investigate the reason(s) why nucleomorph genomes vary in size within cryptomonads and chlorarachniophytes and to speculate why these organelles persist.

Nucleomorph Genomes are "Jam-Packed"

As of May 2011, nucleomorph genome sequences have been published for three cryptomonads, *Guillardia theta* (Douglas et al. 2001), *Hemiselmis andersenii* (Lane et al. 2007), and *Cryptomonas paramecium* (Tanifuji et al. 2011), and for the

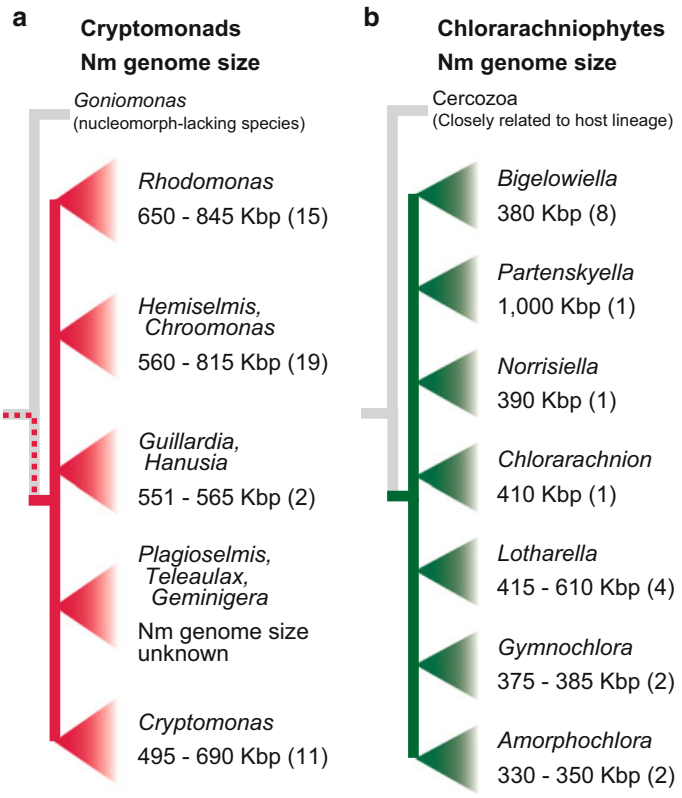


Fig. 2 Nucleomorph genome size variation in cryptomonads (**a**) and chlorarachniophytes (**b**). *Numbers* beside genome sizes show the number of strains examined thus far in each clade. The schematic trees and genome size estimates were taken from Moore and Archibald (2009) with modifications from Ota et al. (2007), Ishida et al. (2011a, b), Tanifuji et al. (2010), and Tanifuji (2011)

chlorarachniophyte *Bigelowiella natans* (Gilson et al. 2006). The salient features of these genomes are shown in Table 1. G + C content of nucleomorph genomes is 25–35 %. This extremely low G + C content is a feature seen in other reduced genomes, such as those of mitochondria, (algal) plastids, and obligate bacterial endosymbionts (Gray et al. 2004; Moran 1996; Nakabachi et al. 2006; Smith 2009). As mentioned above, nucleomorph genomes are characterized by the presence of rDNA operons on their three linear chromosome ends. In cryptomonads, some chromosome ends posses only 5S rDNA, and the direction of the rDNA operons are not always the same in chlorarachniophytes (Douglas et al. 2001; Gilson et al. 2006; Lane et al. 2007; Lane and Archibald 2006, 2008; Silver et al. 2007, 2010; Tanifuji et al. 2011).

Nucleomorph genomes are extremely gene dense: average intergenic spacer length is 94–132 bp and gene density is 0.93–1.10 kbp/gene. Lane et al. (2007)

Table 1 Overview of published nucleomorph genome sequences

Species	*Guillardia theta*	*Hemiselmis andersenii*	*Cryptomonas paramecium*	*Bigelowiella natans*
Genome size (kbp)	Total 550.5 chr.1 195.9 chr.2 180.6 chr.3 173.9	Total 571.4 chr.1 207.3 chr.2 184.6 chr.3 179.4	Total 485.9 chr.1 177.0 chr.2 159.7 chr.3 149.1	Total 372.9 chr.1 140.6 chr.2 134.1 chr.3 98.1
G + C content (%)	26.43	25.18	26.05	<35
Number of genes (protein-coding genes/Total)	487 (548)	472 (525)	466 (519)	293 (340)
Mean polypeptide length (amino acids)	311.66	338.41	289.39	321.90
Mean intergenic distance (base pairs)	94.89	132.14	103.49	113
Gene density (kilobase pairs/gene)	1.00	1.09	0.93	1.10
Number of overlapping genes	44	11	33	Numerous
Number of introns	17	0	2	852
Number of plastid-associated genes	30	31	18	17
Telomere sequences	$(AG)_7AAG_6A$	GA_{17}	GA_9	TCTAGGG

Data taken from Douglas et al. (2001), Gilson et al. (2006), Lane et al. (2007), and Tanifuji et al. (2011). Numbers may vary slightly, depending on updated analyses

compared the 550 kbp nucleomorph genome of *G. theta* to the 571 kbp genome of *H. andersenii*. The *G. theta* genome was found to encode significantly smaller proteins and to be more compact. Moreover, both nucleomorph genomes are more compact compared to those of free-living algae, and the average length of their proteins is also shorter. These results suggest that the process of genome reduction and compaction can impact both coding and noncoding regions of nucleomorph genomes. More recently, however, the 486 kbp genome of the non-photosynthetic species *C. paramecium* was found to possess longer intergenic spacers than *G. theta*, which has a bigger genome (Tanifuji et al. 2011). Therefore, the precise relationship between intergenic distance and mean protein length in nucleomorph genome size evolution is not clear. More nucleomorph genome sequence data, especially from closely related strains and species, are needed to better understand the factors driving nucleomorph genome size change, as well as the relationship between gene density and gene/protein size.

An interesting difference between cryptomonad and chlorarachniophyte nucleomorph genomes is the size and abundance of their spliceosomal introns. While spliceosomal introns are rare in cryptomonad nucleomorph genomes (17 in *G. theta*, 0 in *H. andersenii*, and 2 in *C. paramecium*) and are between 42 and 100 bp in size, introns are abundant in the nucleomorph genome of the chlorarachniophyte *B. natans*, which has 852 introns 18–21 bp in size as well as numerous splicing machinery-related genes (Douglas et al. 2001; Gilson et al. 2006; Lane et al. 2007; Tanifuji et al. 2011). The extent to which these numbers reflect differences in the intron densities of the algal progenitors of the cryptomonad and

chlorarachniophyte nucleomorphs is unclear. The absence of introns and genes for splicing RNAs and proteins in the *H. andersenii* nucleomorph genome was the first instance of complete intron loss in a nuclear genome (Lane et al. 2007).

Another interesting difference between cryptomonad and chlorarachniophyte nucleomorphs is their telomere structures. Within cryptomonads, telomeric repeats differ in primary sequence from species to species, whereas the chlorarachniophytes investigated thus far have identical or very similar sequences (Douglas et al. 2001; Gilson et al. 2006; Lane et al. 2007; Silver et al. 2007, 2010; Tanifuji et al. 2011). As well, the telomere sequences in cryptomonad nucleomorph genomes show no obvious similarity to those of red algae, whereas the chlorarachniophyte telomeres are clearly green algal in nature. The apparent rapid evolution of telomere sequences in cryptomonad nucleomorph genomes is intriguing given that such sequences are typically conserved within evolutionarily diverse lineages such as animals and plants (Podlevsky et al. 2008).

Nucleomorph Gene Content

A fundamental question in plastid evolution is why do nucleomorphs persist in cryptomonads and chlorarachniophytes but have disappeared in all other secondary plastid-bearing organisms. Before the availability of genome sequences, it was predicted that the main function of the nucleomorph was to service the plastid with which it was inherited. Therefore, it was predicted that nucleomorph genomes would encode a plethora of plastid-associated genes that had not yet been transferred to the secondary host nucleus. This proved not to be the case. Douglas et al. (2001) published the complete nucleomorph genome sequence of the cryptomonad *G. theta* in 2001. The genome was found to contain 548 predicted genes, 487 of which code for protein (based on the latest information in GenBank; Table 1), but only 30 of these encode plastid-targeted proteins. The bulk of the *G. theta* nucleomorph proteins are in fact "housekeeping" proteins involved in processes such as translation, transcription, and nucleic acid metabolism in the PPC. More recently, precisely the same set of 30 plastid-associated genes [31 genes including the analyses of Tanifuji et al. (2011)] were found amongst 472 protein coding genes in *H. andersenii* (Lane et al. 2007). The secondarily non-photosynthetic cryptomonad *C. paramecium* retains 18 plastid-associated genes (out of 466 predicted protein coding genes in total) (Tanifuji et al. 2011). The nucleomorph genome of the chlorarachniophyte *B. natans* possesses only 17 plastid-associated genes out of 293 protein genes in total (Gilson et al. 2006). These 17 genes do not overlap with those found in the cryptomonads. Thus, only 3–6 % of the genes in known nucleomorph genomes appear to be associated with the plastid.

Based on the complete lack of overlap between the plastid-associated genes found in sequenced cryptomonad and chlorarachniophyte genomes, Gilson et al. (2006) speculated that nucleomorphs and their genomes could completely

disappear if the residual set of plastid protein genes were transferred to the secondary host nucleus. Based on analysis of an expanded dataset, Tanifuji et al. (2011) suggested that there is indeed no obvious reason why these plastid-associated genes persist in present-day nucleomorph genomes on the following grounds. First, the genome of the cryptomonad *C. paramecium* has fewer plastid-associated genes than do the *H. andersenii* and *G. theta* nucleomorphs, including those encoding proteins that are plastid-targeted but *not* involved in photosynthesis. Second, eight of 13 plastid-associated genes shared among the three cryptomonad nucleomorphs reside in the host nucleus of the haptophyte *Emiliania huxleyi*. The haptophytes have been suggested to be specifically related to cryptomonads from the perspective of the secondary host [e.g., Burki et al. (2008) and Patron et al. (2007)]. However, as we shall see, key insight into the reasons why nucleomorphs persist in cryptomonads and chlorarachniophytes has come from considering a mysterious class of nucleomorph genes for which sequence similarity comparisons provide no insight into their persistence.

ORFan Genes in Nucleomorph Genomes

Approximately, 65 % of nucleomorph genes (including structural RNA genes) have obvious homologs (or conserved functional domains) in other eukaryotes. As noted above, most of these are housekeeping in nature. However, the remaining 35 % are not only genes of unknown function, but they also either show no primary sequence similarity whatsoever to any other sequences in other published gene/genome databases or have a detectable homolog only in other nucleomorph genomes. Tanifuji et al. (2011) designated these genes as nucleomorph ORFans, or "nORFans," and investigated how many nORFans are shared among cryptomonad nucleomorphs. Only ~12 % of the nORFans (23 genes in total) are conserved amongst the *G. theta, H. andersenii*, and *C. paramecium* nucleomorph genomes (Fig. 3a). These numbers are in stark contrast to the ~90 % (213 genes) of protein-coding housekeeping genes shared between these three genomes (Fig. 3a). 70–81 % of cryptomonad nORFans have no obvious homolog in another nucleomorph genome (or by definition in any other genome), accounting for 24–29 % of the total number of cryptomonad nucleomorph genes. In general, ORFans are rare in reduced genomes, the rationale being that only essential, highly conserved genes for core cellular processes are retained, while the "species-specific" or variable genes should disappear. For reference, only six ORFans are found in the plastid genomes of the cryptomonads *C. paramecium*, *R. salina*, and *G. theta* (Douglas and Penny 1999; Donaher et al. 2009; Khan et al. 2007). In *Buchnera* sp. APS, an obligate endosymbiont of insects, only 7 ORFans were found among its 575 -protein-coding genes (Degnan et al. 2005; Moran et al. 2008; Shigenobu et al. 2000).

Interestingly, many nORFans have counterparts in syntenic regions in other nucleomorph genomes, despite the presence of no discernable sequence similarity.

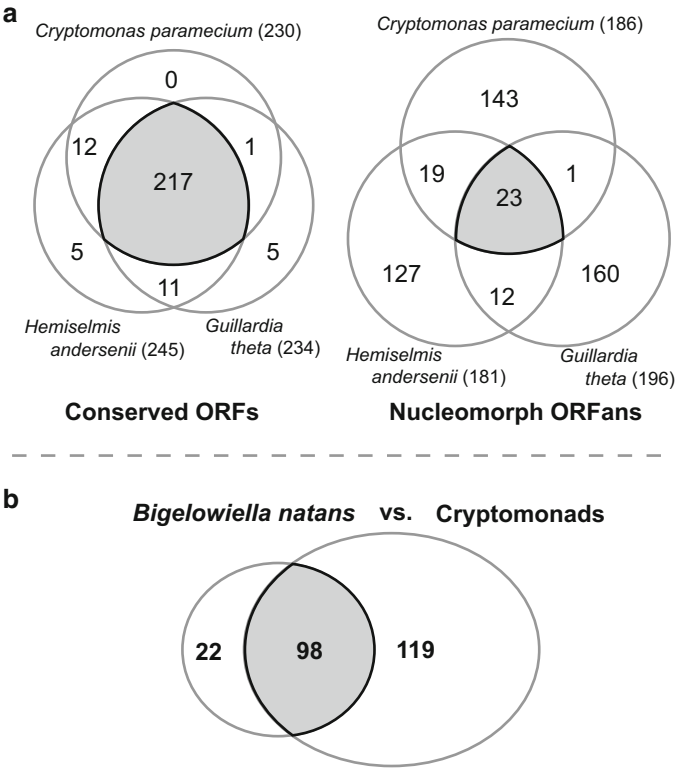


Fig. 3 Nucleomorph genome content variation. (**a**) Comparison of *Guillardia theta*, *Hemiselmis andersenii*, and *Cryptomonas paramecium* nucleomorph genes. Venn diagrams show the number of shared and/or unique genes when conserved ORFs (i.e., those with obvious homologs, or homologous domains, in other eukaryotes) and nucleomorph ORFans (i.e., protein-coding genes found only in nucleomorph genomes; see main text) are compared, respectively. (**b**) Comparison of the 217 shared conserved ORFs in the three cryptomonad nucleomorph genomes and housekeeping genes in the *B. natans* nucleomorph. Modified from Tanifuji et al. (2011)

Figure 4 shows an example of the phenomenon of "syntenic nORFans." Many of these genes are embedded within blocks of conserved genes in the same order and direction in all three sequenced cryptomonad nucleomorph genomes and are typically similar in size. This suggests that the syntenic nORFans are homologous despite showing no obvious similarity to one another. Why are they retained? Given that many nORFans encode proteins comprised of amino acids corresponding to A + T-rich codons, some of which are basic, hydrophobic, and polar residues (phenylalanine, isoleucine, asparagine, lysine, and tyrosine), Archibald and Lane (2009) suggested that nORFans could encode membrane interacting/transmembrane proteins whose primary amino acid sequence is not constrained (Archibald and Lane 2009; Deber et al. 1999). This hypothesis should be tested in the future, but for now the functional significance of nORFans is an unsolved mystery in nucleomorph research.

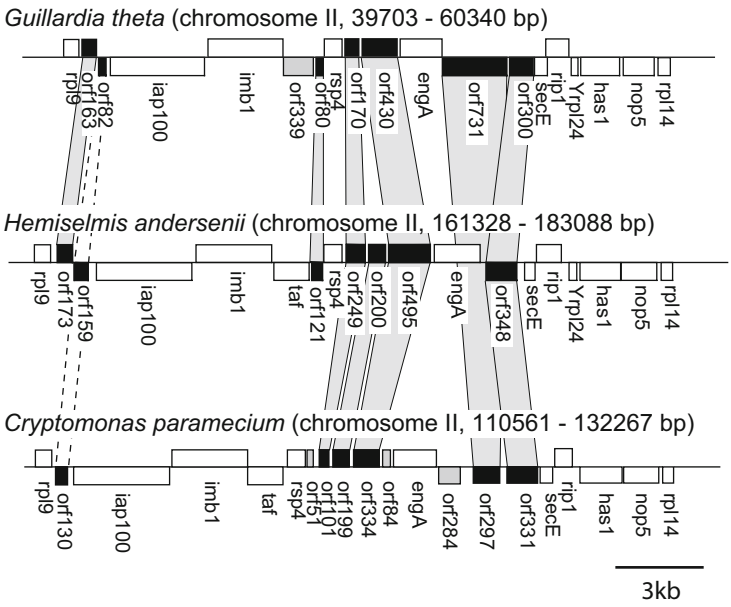


Fig. 4 Example of gene order conservation (synteny) in the nucleomorph genomes of the cryptomonads *Guillardia theta*, *Hemiselmis andersenii*, and *Cryptomonas paramecium*. *Black and gray boxes* show nucleomorph ORFans (nORFans). *Black boxes* indicate syntenic nORFans, and *gray shadows* emphasize their conserved position within the same syntenic block. *Dotted lines* show syntenic nORFans with detectable amino acid sequence similarity. Genes indicated by *boxes above the lines* are transcribed left to right, those *below the lines* right to left. Data taken from Douglas et al. (2001), Lane et al. (2007), and Tanifuji et al. (2011)

Nucleomorphs: End Result or Intermediate State?

The three-chromosome architecture of cryptomonad and chlorarachniophyte nucleomorphs represents a striking example of convergent genome evolution. To what extent is this convergence apparent at the level of genes? In an initial two-way comparison of genes in the nucleomorph of the cryptomonad *G. theta* and the chlorarachniophyte *B. natans*, Gilson et al. (2006) concluded that while both genomes were "enriched" in housekeeping genes, there was very little overlap between the two on a gene-by-gene basis. Tanifuji et al. (2011) revisited this question by comparing the 217 genes shared between three cryptomonad nucleomorph genomes with 120 protein genes in the *B. natans* nucleomorph genome (splicing-related genes were excluded, due to the fact that the *H. andersenii* genome lacks introns). The results were unexpected: 81 % of the analyzable protein genes in the *B. natans* genome (i.e., 98/120 genes) are also found in the "core" cryptomonad nucleomorph gene set (Fig. 3b). For example, in the functional category of translation, 49 of 58 *B. natans* proteins are also found in the cryptomonad nucleomorph genomes. Similar patterns are also seen for transcription

and protein folding/degradation (Tanifuji et al. 2011). Rather than a pattern of "random" gene retention, these observations raise the intriguing possibility that the cryptomonads and chlorarachniophytes maintain their nucleomorphs for similar reasons.

So are nucleomorph genomes in an intermediate state of reductive evolution or have they reached an endpoint? There is presently not enough data to answer this question. As mentioned above, no simple, compelling reason has thus far been put forth to explain why cryptomonad and chlorarachniophyte nucleomorphs should persist, yet there are numerous secondary plastid-bearing lineages where nucleomorphs have disappeared (Gould et al. 2008; Reyes-Prieto et al. 2007; Gagat et al. 2013, pp. xx–yy). What could explain these observations? In order for an essential nucleomorph gene to be lost, the host nucleus must "acquire" a suitable homolog capable of being targeted back to the endosymbiont compartment prior to or concomitant with nucleomorph gene loss. The following scenarios are possible: (i) the nucleomorph gene could be transferred to the host nucleus by EGT and acquire the ability to encode a protein with the necessary targeting information; (ii) gene duplication could lead to a situation in which a host-derived homolog evolves the ability to target its gene product as a substitute for a nucleomorph counterpart; (iii) a single host-derived locus could play the same role for both the host and endosymbiont, possibly through the production of proteins with different amino termini by alternative splicing; or (iv) a functional replacement could be acquired by lateral gene transfer, initially with or without the ability to encode a protein with suitable targeting information. The full extent to which some or all of these processes have contributed to the process of nucleomorph gene elimination is at present unclear. One potentially important fact is that analyses of the host nuclear genomes of *G. theta* and *B. natans* indicate that nucleomorph-to-host-nucleus gene transfer in *G. theta* and *B. natans* appears to be extremely rare (Curtis et al. 2012). This could mean that the window of time during which EGT initially took place in cryptomonads and chlorarachniophytes might have already passed, limiting the opportunity for further gene loss and nucleomorph genome reduction. In this sense, their nucleomorph genomes could be considered "frozen," in the same way as the mitochondrial genomes of animals (Daley and Whelan 2005; Saccone et al. 2002).

Detailed examination of the subcellular proteomes of *G. theta* and *B. natans* (Curtis et al. 2012) provided additional insight into the evolution of nucleomorphs. For example, nucleus-encoded, nucleomorph/PPC-targeted phosphoglucomutase, starch synthase class I-II, and cyclin-dependent kinase regulatory subunit proteins appear not to be of nucleomorph origin but rather to have evolved from a duplicated host gene (scenario ii above). Putative cases of differential targeting of protein isoforms by alternative splicing and alternative translation start site selection were also found in *B. natans* (scenario iii) [see also Hirakawa et al. (2012)], and the evolutionary origins of PPC-targeted proteins in both *G. theta* and *B. natans* were found to be unexpectedly diverse (consistent with a possible role for lateral gene transfer; scenario iv). Curtis et al. (2012) considered the extent to which the PPC proteomes of *G. theta* and *B. natans* appear to be reduced compared to free-living

organisms, and consequently, how core cellular processes might be impacted. The total number of predicted PPC proteins in *B. natans* (1,002) was far fewer than that of *G. theta* (2,401), and interesting differences were seen upon examination of specific biochemical subsystems. For example, 17 of 79 "core" ribosomal protein subunits were found to be missing from the PPC of *B. natans* (taking into account both nucleomorph- and nucleus-encoded, PPC-targeted proteins), more than the predicted number of missing proteins in *G. theta*. In contrast, *B. natans* has a larger set of nucleomorph-targeted proteins involved in splicing than does *G. theta*. In sum, taking into account the coding capacities of present-day nucleomorph genomes and nucleus-encoded, nucleomorph-/PPC-targeted proteins, it is presently still not possible to provide a single, complete answer to the question of why nucleomorphs persist.

Conclusions and Future Directions

Secondary endosymbiosis has been a major factor driving the spread of photosynthesis in eukaryotes. The cryptomonads and chlorarachniophytes are interesting organisms in which to study this process, given that they represent a midpoint in the transition from secondary endosymbiont to nucleomorph-lacking secondary plastid, as seen in haptophytes, heterokonts, and other algae. Despite their independent origins, cryptomonad and chlorarachniophyte nucleomorphs share intriguing similarities in genome size, structure, and even gene content. Although nucleomorph genome sequences have provided some insight into the nature of the processes underlying their evolution, many questions still remain. Why do nucleomorph genomes contain three chromosomes? Why are rDNA operons almost invariably located near the termini of nucleomorph chromosomes? What are the functions of nORFans? The central mysteries of nucleomorph biology are also still unresolved, e.g., why do they persist only in cryptomonads and chlorarachniophytes? The recently published nuclear genome sequences from the cryptomonad *G. theta* and the chlorarachniophyte *B. natans* are an important first step to answering these and related questions, but are only a start point. Such data will become much more useful when compared to nucleomorph and nuclear genome data from other species within both lineages. Given the pace and cost at which genome sequences can now be obtained, the data should soon be forthcoming.

Acknowledgments We thank Dr. Shuhei Ota for providing light micrographs and TEM pictures of chlorarachniophytes and Christa Moore for comments on an earlier version of this manuscript. Nucleomorph genome sequencing projects in the Archibald Laboratory are supported by an operating grant from the Canadian Institutes of Health Research (CIHR) Regional Partnership Program, together with the Nova Scotia Health Research Foundation. GT is supported by a Tula Foundation postdoctoral fellowship from the Centre for Comparative Genomics and Evolutionary Bioinformatics at Dalhousie University. JMA is a Fellow of the Canadian Institute for Advanced Research, Program in Integrated Microbial Biodiversity, and holder of a New Investigator Award from the Canadian Institutes of Health Research.

References

Archibald JM (2007) Nucleomorph genomes: structure, function, origin and evolution. Bioessays 29:392–402

Archibald JM, Lane CE (2009) Going, going, not quite gone: nucleomorphs as a case study in nuclear genome reduction. J Hered 100:582–590

Bock R, Timmis JN (2008) Reconstructing evolution: gene transfer from plastids to the nucleus. Bioessays 30:556–566

Burki F, Shalchian-Tabrizi K, Pawlowski J (2008) Phylogenomics reveals a new 'megagroup' including most photosynthetic eukaryotes. Bio Lett 4:366–369

Burki F, Inagaki Y, Brate J, Archibald JM, Keeling PJ, Cavalier-Smith T, Sakaguchi M, Hashimoto T, Horak A, Kumar S, Klaveness D, Jakobsen KS, Pawlowski J, Shalchian-Tabrizi K (2009) Large-scale phylogenomic analyses reveal that two enigmatic protist lineages, telonemia and centroheliozoa, are related to photosynthetic chromalveolates. Genome Biol Evol 1:231–238

Cavalier-Smith T (2002) Nucleomorphs: enslaved algal nuclei. Curr Opin Microbiol 5:612–619

Cavalier-Smith T, Couch JA, Thorsteinsen KE, Gilson P, Deane JA, Hill DRA, McFadden GI (1996) Cryptomonad nuclear and nucleomorph 18S rRNA phylogeny. Eur J Phycol 31:315–328

Curtis BA, Tanifuji G, Burki F, Gruber A, Irimia M, Maruyama S, Arias MC, Ball SG, Gile GH, Hirakawa Y, Hopkins JF, Kuo A, Rensing SA, Schmutz J, Symeonidi A, Elias M, Eveleigh RJM, Herman EK, Klute MJ, Nakayama T, Oborník M, Reyes-Prieto A, Armbrust EV, Aves SJ, Beiko RG, Coutinho P, Dacks JB, Durnford DG, Fast NM, Green BR, Grisdale C, Hempel F, Henrissat B, Höppner MP, Ishida K, Kim E, Kořený L, Kroth PG, Liu Y, Malik S-B, Maier UG, McRose D, Mock T, Neilson JAD, Onodera NT, Poole AM, Pritham EJ, Richards TA, Rocap G, Roy SW, Sarai C, Schaack S, Shirato S, Slamovits CH, Spencer DF, Suzuki S, Worden AZ, Zauner S, Barry K, Bell C, Bharti AK, Crow JA, Grimwood J, Kramer R, Lindquist E, Lucas S, Salamov A, McFadden GI, Lane CE, Keeling PJ, Gray MW, Grigoriev IV, Archibald JM (2012) Algal genomes reveal evolutionary mosaicism and the fate of nucleomorphs. Nature 492:59–65

Daley DO, Whelan J (2005) Why genes persist in organelle genomes. Genome Biol 6:110

Deane JA, Isabbelle M, Strachan IM, Saunders GW, Hill DRA, McFadden GI (2002) Cryptomonad evolution: nuclear 18SrDNA phycologeny versus cell morphology and pigmentation. J Phycol 38:1236–1244

Deber CM, Liu LP, Wang C (1999) Perspective: peptides as mimics of transmembrane segments in proteins. J Pept Res 54:200–205

Degnan PH, Lazarus AB, Wernegreen JJ (2005) Genome sequence of *Blochmannia pennsylvanicus* indicates parallel evolutionary trends among bacterial mutualists of insects. Genome Res 15:1023–1033

Derelle E, Ferraz C, Rombauts S, Rouze P, Worden AZ, Robbens S, Partensky F, Degroeve S, Echeynie S, Cooke R, Saeys Y, Wuyts J, Jabbari K, Bowler C, Panaud O, Piegu B, Ball SG, Ral JP, Bouget FY, Piganeau G, De Baets B, Picard A, Delseny M, Demaille J, Van de Peer Y, Moreau H (2006) Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. Proc Natl Acad Sci U S A 103:11647–11652

Dolezal P, Likic V, Tachezy J, Lithgow T (2006) Evolution of the molecular machines for protein import into mitochondria. Science 313:314–318

Donaher N, Tanifuji G, Onodera NT, Malfatti SA, Chain PSG, Hara Y, Archibald JM (2009) The complete plastid genome sequence of the secondarily nonphotosynthetic alga *Cryptomonas paramecium*: reduction, compaction, and accelerated evolutionary rate. Genome Biol Evol 1:439–448

Douglas SE, Penny SL (1999) The plastid genome of the cryptophyte alga, *Guillardia theta*: complete sequence and conserved synteny groups confirm its common ancestry with red algae. J Mol Evol 48:236–244

Douglas SE, Durnford DG, Morden CW (1990) Nucleotide sequence of the gene for the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase from *Cryptomonas phi* : evidence supporting the polyphyletic origin of plastids. J Phycol 26:500–508

Douglas S, Zauner S, Fraunholz M, Beaton M, Penny S, Deng LT, Wu XN, Reith M, Cavalier-Smith T, Maier UG (2001) The highly reduced genome of an enslaved algal nucleus. Nature 410:1091–1096

Eschbach S, Hofmann CJB, Maier UG, Sitte P, Hansmann P (1991) A eukaryotic genome of 660 Kb: electrophoretic karyotype of nucleomorph and cell nucleus of the cryptomonad alga, *Pyrenomonas salina*. Nucleic Acids Res 19:1779–1781

Gagat P, Mackiewicz P, Bodył A (2013) Secondary endosymbioses. In: Löffelhardt W (ed) Endosymbiosis. Springer, Vienna, pp xx–yy

Gibbs SP (1978) Chloroplasts of euglena may have evolved from symbiotic green-algae. Can J Bot 56:2883–2889

Gilson PR, McFadden GI (2002) Jam packed genomes – a preliminary, comparative analysis of nucleomorphs. Genetica 115:13–28

Gilson PR, Su V, Slamovits CH, Reith ME, Keeling PJ, McFadden GI (2006) Complete nucleotide sequence of the chlorarachniophyte nucleomorph: nature's smallest nucleus. Proc Natl Acad Sci U S A 103:9566–9571

Gould SB, Waller RR, McFadden GI (2008) Plastid evolution. Annu Rev Plant Biol 59:491–517

Gray MW, Doolittle WF (1982) Has the endosymbiont hypothesis been proven? Microbiol Rev 46:1–42

Gray MW, Burger G, Lang BF (1999) Mitochondrial evolution. Science 283:1476–1481

Gray MW, Lang BF, Burger G (2004) Mitochondria of protists. Annu Rev Genet 38:477–524

Greenwood A (1974) The cryptophyta in relation to phylogeny and photosynthesis. In: Proceedings of the 8th International Congress on Electron Microscopy, vol 2, pp 566–567

Hackett JD, Anderson DM, Erdner DL, Bhattacharya D (2004) Dinoflagellates: a remarkable evolutionary experiment. Am J Bot 91:1523–1534

Hampl V, Hug L, Leigh JW, Dacks JB, Lang BF, Simpson AGB, Roger AJ (2009) Phylogenomic analyses support the monophyly of excavata and resolve relationships among eukaryotic "supergroups". Proc Natl Acad Sci U S A 106:3859–3864

Hibberd DJ, Norris RE (1984) Cytology and ultrastructure of *Chlorarachnion reptans* (Chlorarachniophyta divisio nova, Chlorarachniophyceae classis nova). J Phycol 20:310–330

Hirakawa Y, Burki F, Keeling PJ (2012) Dual targeting of aminoacyl-tRNA synthetases to the mitochondrion and plastid in chlorarachniophytes. J Cell Sci 125(Pt 24):6176–6184

Hoef-Emden K (2005) Multiple independent losses of photosynthesis and differing evolutionary rates in the genus *Cryptomonas* (cryptophyceae): combined phylogenetic analyses of DNA sequences of the nuclear and the nucleomorph ribosomal operons. J Mol Evol 60:183–195

Hoef-Emden K, Marin B, Melkonian M (2002) Nuclear and nucleomorph SSU rDNA phylogeny in the cryptophyta and the evolution of cryptophyte diversity. J Mol Evol 55:161–179

Ishida K, Cao Y, Hasegawa M, Okada N, Hara Y (1997) The origin of chlorarachniophyte plastids, as inferred from phylogenetic comparisons of amino acid sequences of EF-Tu. J Mol Evol 45:682–687

Ishida K, Green BR, Cavalier-Smith T (1999) Diversification of a chimaeric algal group, the chlorarachniophytes: phylogeny of nuclear and nucleomorph small-subunit rRNA genes. Mol Biol Evol 16:321–331

Ishida K, Endo H, Koike S (2011a) *Partenskyella glossopodia* (Chlorarachniophyceae) possesses a nucleomorph genome of approximately 1 Mbp. Phycol Res 59:120–122

Ishida K, Yabuki A, Ota S (2011b) *Amorphochlora amoebiformis* gen. et comb. nov. (Chlorarachniophyceae). Phycol Res 59:52–53

Katinka MD, Duprat S, Cornillot E, Metenier G, Thomarat F, Prensier G, Barbe V, Peyretaillade E, Brottier P, Wincker P, Delbac F, El Alaoui H, Peyret P, Saurin W, Gouy M, Weissenbach J, Vivares CP (2001) Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. Nature 414:450–453

Khan H, Parks N, Kozera C, Curtis BA, Parsons BJ, Bowman S, Archibald JM (2007) Plastid genome sequence of the cryptophyte alga *Rhodomonas salina* CCMP1319: lateral transfer of putative DNA replication machinery and a test of chromist plastid phylogeny. Mol Biol Evol 24:1832–1842

Lane CE, Archibald JM (2006) Novel nucleomorph genome architecture in the cryptomonad genus *Hemiselmis*. J Eukaryot Microbiol 53:515–521

Lane CE, Archibald JM (2008) New marine members of the genus *Hemiselmis* (Cryptomonadales, Cryptophyceae). J Phycol 44:439–450

Lane CE, Khan H, MacKinnon M, Fong A, Theophilou S, Archibald JM (2006) Insight into the diversity and evolution of the cryptomonad nucleomorph genome. Mol Biol Evol 23:856–865

Lane CE, van den Heuvel K, Kozera C, Curtis BA, Parsons BJ, Bowman S, Archibald JM (2007) Nucleomorph genome of *Hemiselmis andersenii* reveals complete intron loss and compaction as a driver of protein structure and function. Proc Natl Acad Sci U S A 104:19908–19913

Löffelhardt W (2013) The single primary endosymbiotic event. In: Löffelhardt W (ed) Endosymbiosis. Springer, Vienna, pp xx–yy

Ludwig M, Gibbs SP (1985) DNA is present in the nucleomorph of cryptomonads – further evidence that the chloroplast evolved from a eukaryotic endosymbiont. Protoplasma 127:9–20

Martin W (2003) The smoking gun of gene transfer. Nat Genet 33:442

Martin W, Herrmann RG (1998) Gene transfer from organelles to the nucleus: how much, what happens, and why? Plant Physiol 118:9–17

Martin W, Schnarrenberger C (1997) The evolution of the Calvin cycle from prokaryotic to eukaryotic chromosomes: a case study of functional redundancy in ancient pathways through endosymbiosis. Curr Genet 32:1–18

Martin W, Mustafa AZ, Henze K, Schnarrenberger C (1996) Higher-plant chloroplast and cytosolic fructose-1,6-bisphosphatase isoenzymes: origins via duplication rather than prokaryote-eukaryote divergence. Plant Mol Biol 32:485–491

Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, Leister D, Stoebe B, Hasegawa M, Penny D (2002) Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proc Natl Acad Sci U S A 99:12246–12251

Matsuzaki M, Misumi O, Shin-I T, Maruyama S, Takahara M, Miyagishima SY, Mori T, Nishida K, Yagisawa F, Nishida K, Yoshida Y, Nishimura Y, Nakao S, Kobayashi T, Momoyama Y, Higashiyama T, Minoda A, Sano M, Nomoto H, Oishi K, Hayashi H, Ohta F, Nishizaka S, Haga S, Miura S, Morishita T, Kabeya Y, Terasawa K, Suzuki Y, Ishii Y, Asakawa S, Takano H, Ohta N, Kuroiwa H, Tanaka K, Shimizu N, Sugano S, Sato N, Nozaki H, Ogasawara N, Kohara Y, Kuroiwa T (2004) Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. Nature 428:653–657

McFadden GI, Gilson PR, Hill DRA (1994) *Goniomonas* – ribosomal RNA sequences indicate that this phagotrophic flagellate is a close relative of the host component of cryptomonads. Eur J Phycol 29:29–32

McFadden GI, Gilson PR, Waller RF (1995) Molecular phylogeny of chlorarachniophytes based on plastid ribosomal RNA and *rbcL* sequences. Arch Protistenkd 145:231–239

Moore CE, Archibald JM (2009) Nucleomorph genomes. Annu Rev Genet 43:251–264

Moran NA (1996) Accelerated evolution and Muller's rachet in endosymbiotic bacteria. Proc Natl Acad Sci U S A 93:2873–2878

Moran NA, McCutcheon JP, Nakabachi A (2008) Genomics and evolution of heritable bacterial symbionts. Annu Rev Genet 42:165–190

Moreira D, Lopezarchilla AI, Amils R, Marin I (1994) Characterization of 2 New thermoacidophilic microalgae – genome organization and comparison with *Galdieria sulphuraria*. FEMS Microbiol Lett 122:109–114

Muravenko OV, Selyakh IO, Kononenko NV, Stadnichuk IN (2001) Chromosome numbers and nuclear DNA contents in the red microalgae *Cyanidium caldarium* and three *Galdieria* species. Eur J Phycol 36:227–232

Nakabachi A, Yamashita A, Toh H, Ishikawa H, Dunbar HE, Moran NA, Hattori M (2006) The 160-kilobase genome of the bacterial endosymbiont *Carsonella*. Science 314:267

Ota S, Ueda K, Ishida K (2007) *Norrisiella sphaerica* gen. et sp nov., a new coccoid chlorarachniophyte from Baja California, Mexico. J Plant Res 120:661–670

Patron NJ, Inagaki Y, Keeling PJ (2007) Multiple gene phylogenies support the monophyly of cryptomonad and haptophyte host lineages. Curr Biol 17:887–891

Phipps KD, Donaher NA, Lane CE, Archibald JM (2008) Nucleomorph karyotype diversity in the freshwater cryptophyte genus *Cryptomonas*. J Phycol 44:11–14

Podlevsky JD, Bley CJ, Omana RV, Qi XD, Chen JJL (2008) The telomerase database. Nucleic Acids Res 36:D339–D343

Rensing SA, Goddemeier M, Hofmann CJB, Maier UG (1994) The presence of a nucleomorph hsp70 gene is a common feature of cryptophyta and chlorarachniophyta. Curr Genet 26:451–455

Reyes-Prieto A, Weber APM, Bhattacharya D (2007) The origin and establishment of the plastid in algae and plants. Annu Rev Genet 41:147–168

Saccone C, Gissi C, Reyes A, Larizza A, Sbisa E, Pesole G (2002) Mitochondrial DNA in metazoa: degree of freedom in a frozen event. Gene 286:3–12

Shalchian-Tabrizi K, Brate J, Logares R, Klaveness D, Berney C, Jakobsen KS (2008) Diversification of unicellular eukaryotes: cryptomonad colonizations of marine and fresh waters inferred from revised 18S rRNA phylogeny. Environ Microbiol 10:2635–2644

Shigenobu S, Watanabe H, Hattori M, Sakaki Y, Ishikawa H (2000) Genome sequence of the endocellular bacterial symbiont of aphids *Buchnera* sp. APS. Nature 407:81–86

Silver TD, Koike S, Yabuki A, Kofuji R, Archibald JM, Ishida KI (2007) Phylogeny and nucleomorph karyotype diversity of chlorarachniophyte algae. J Eukaryot Microbiol 54:403–410

Silver TD, Moore CE, Archibald JM (2010) Nucleomorph ribosomal DNA and telomere dynamics in chlorarachniophyte algae. J Eukaryot Microbiol 57:453–459

Smith DR (2009) Unparalleled GC content in the plastid DNA of *Selaginella*. Plant Mol Biol 71:627–639

Tanifuji G (2011) Reductive genome evolution in nucleomorphs (in Japanese). Jpn J Protozool 44:89–102

Tanifuji G, Onodera NT, Hara Y (2010) Nucleomorph genome diversity and its phylogenetic implications in cryptomonad algae. Phycol Res 58:230–237

Tanifuji G, Onodera NT, Wheeler TJ, Dlutek M, Donaher N, Archibald JM (2011) Complete nucleomorph genome sequence of the nonphotosynthetic alga *Cryptomonas paramecium* reveals a core nucleomorph gene set. Genome Biol Evol 3:44–54

van Dooren GG, Schwartzbach SD, Osafune T, McFadden GI (2001) Translocation of proteins across the multiple membranes of complex plastids. Biochim Biophys Acta 1541:34–53

Whatley JM, Whatley FR (1981) Chloroplast evolution. New Phytol 87:233–247

Worden AZ, Lee JH, Mock T, Rouze P, Simmons MP, Aerts AL, Allen AE, Cuvelier ML, Derelle E, Everett MV, Foulon E, Grimwood J, Gundlach H, Henrissat B, Napoli C, McDonald SM, Parker MS, Rombauts S, Salamov A, Von Dassow P, Badger JH, Coutinho PM, Demir E, Dubchak I, Gentemann C, Eikrem W, Gready JE, John U, Lanier W, Lindquist EA, Lucas S, Mayer KFX, Moreau H, Not F, Otillar R, Panaud O, Pangilinan J, Paulsen I, Piegu B, Poliakov A, Robbens S, Schmutz J, Toulza E, Wyss T, Zelensky A, Zhou K, Armbrust EV, Bhattacharya D, Goodenough UW, Van de Peer Y, Grigoriev IV (2009) Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes *Micromonas*. Science 324:268–272

Protein Import into Complex Plastids: Current Findings and Perspectives

Christopher Grosche, Franziska Hempel, Kathrin Bolte, Lars Abram, Uwe G. Maier, and Stefan Zauner

Abstract Transport of proteins across either two, three, four, or even five membranes is a feature of plastids, which evolved via the engulfment of a phototrophic eukaryote by another eukaryotic cell (secondary/tertiary endosymbiosis). Although emerging data are helpful for a mechanistic explanation of protein transport across the membranes surrounding secondary plastids, several questions still have to be answered. Here, we describe the recent models concerning protein import into secondary plastids and discuss their implications.

Contents

Introduction 216
Preprotein Transport Across the Membranes of Secondary Plastids: A Brief Summary of the Status Quo 216
Open Questions on Preprotein Transport into Complex Plastids 220
Crossing the Outermost Membrane 220
Integration of Proteins into the Outermost Membrane 221
The Role of the N-Terminal Targeting Signals 222

C. Grosche • F. Hempel • U.G. Maier (✉)
Laboratorium für Zellbiologie, Philipps-Universität Marburg, Karl-von-Frisch Strasse 8, 35032 Marburg, Germany

LOEWE-Zentrum für Synthetische Mikrobiologie (SynMikro), Hans-Meerwein-Strasse, 35032 Marburg, Germany
e-mail: maier@staff.uni-marburg.de

K. Bolte • S. Zauner
Laboratorium für Zellbiologie, Philipps-Universität Marburg, Karl-von-Frisch Strasse 8, 35032 Marburg, Germany

L. Abram
Laboratorium für Zellbiologie, Philipps-Universität Marburg, Karl-von-Frisch Strasse 8, 35032 Marburg, Germany

Institut für Zell- und Entwicklungsbiologie der Pflanzen, Universität Bielefeld, Universitätsstrasse 25, 33615 Bielefeld, Germany

W. Löffelhardt (ed.), *Endosymbiosis*, DOI 10.1007/978-3-7091-1303-5_12,
© Springer-Verlag Wien 2014

Transport of Membrane Proteins Destined for the Second, Third, or Fourth Membrane 223
PPC Versus Stroma Targeting in Plastids Surrounded by Four Membranes 223
Recycling the Old for Generating the New 225
Conclusions 227
References 227

Introduction

Plastids are important cellular organelles involved not only in photosynthesis but also in many other essential cellular functions. They can be traced back to once free-living prokaryotic cells in case of primary endosymbiosis, or even to eukaryotic, phototrophic cells in case of secondary and tertiary endosymbioses (Keeling 2009). Due to the fact that the plastid proteome is expressed in the organelle only to a minor extent (Martin et al. 2002; Hempel et al. 2007), import of nucleus-encoded plastid proteins is essential to maintain plastid functionality.

In the easiest, but still very complicated case, the host-expressed plastid proteins have to cross the two surrounding membranes of primary plastids, present in green algae, land plants, red algae, and the glaucophytes (Soll and Schleiff 2004; Hempel et al. 2007; Bolte et al. 2009). In most cases, these proteins are expressed within the cytoplasm of the host cell. They are synthesized as preproteins containing a transit peptide, an N-terminal targeting sequence for the import across the plastid envelope (Bionda et al. 2010). Additional plastid membranes have to be traversed in organisms evolved via secondary endosymbiosis. Here, an already phototrophic eukaryotic cell, either a green or a red alga, was engulfed by another eukaryotic cell and reduced to a so-called secondary or complex plastid that is surrounded by additional membranes in comparison to primary plastids (Maier et al. 2000; Gould et al. 2008). Thus, further protein translocation machineries had to evolve to guarantee proper protein transport across these additional membranes.

Preprotein Transport Across the Membranes of Secondary Plastids: A Brief Summary of the Status Quo

The majority of secondarily evolved plastids are surrounded by four membranes as it is known for apicomplexa, heterokontophytes, haptophytes, cryptophytes, and chlorarachniophytes. Except for the last, which harbor a green algal endosymbiont, these organisms contain a complex plastid of red algal origin and are thus representing the red lineage (Cavalier-Smith 2003). Secondary plastids of euglenophytes, which evolved by the uptake of an ancestral green algae, and peridinin-containing dinoflagellates, which in contrast harbor a plastid of red algal origin (Archibald 2009; Janouškovec et al. 2011), are surrounded by only three membranes (Fig. 1).

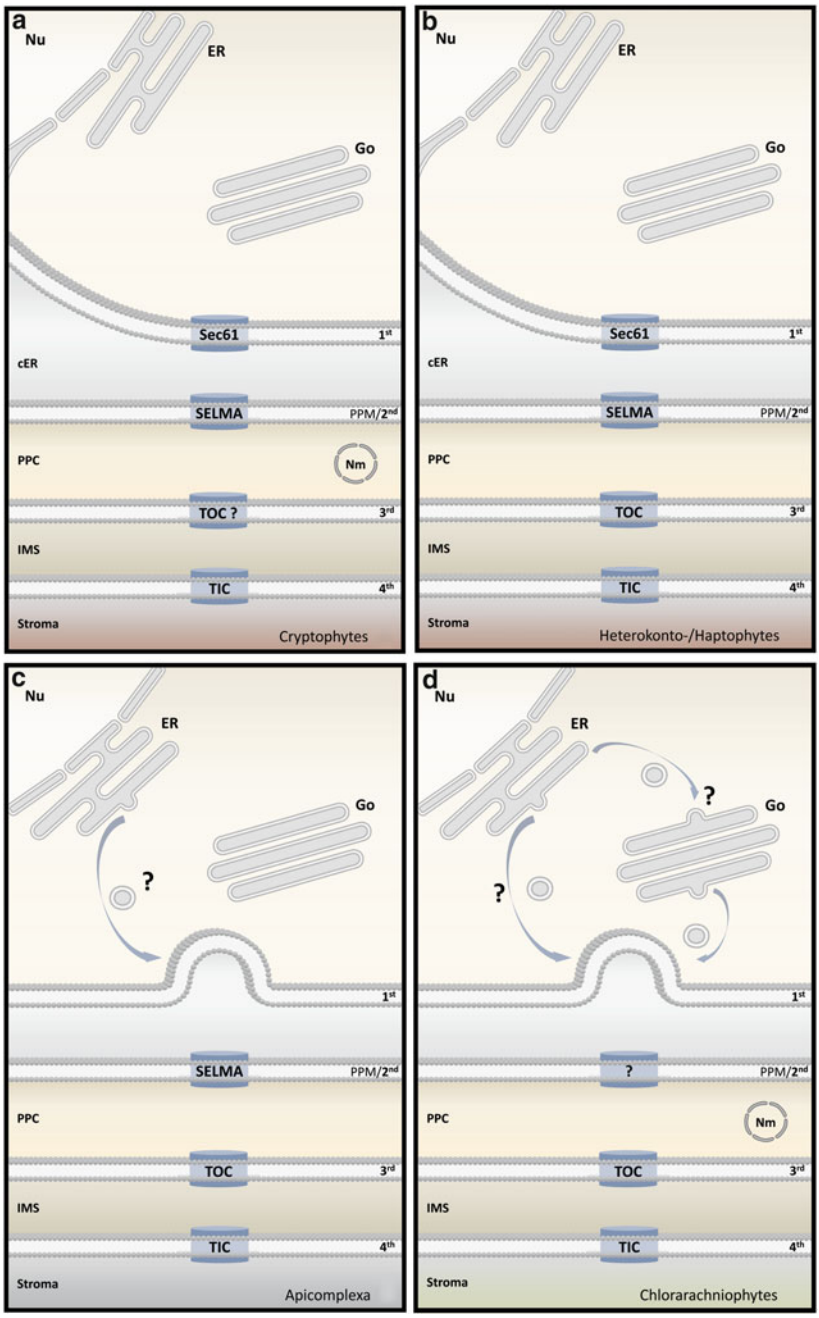


Fig. 1 (continued)

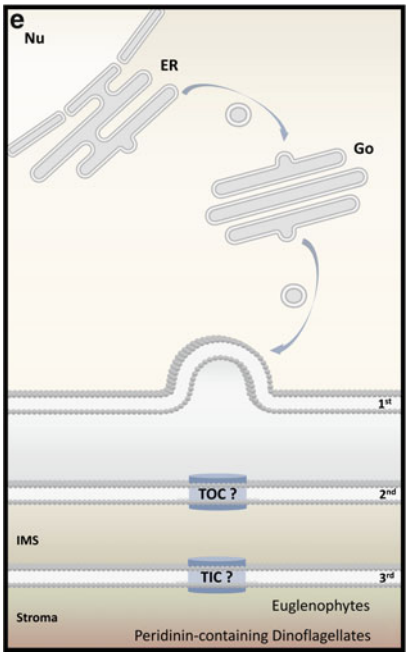


Fig. 1 Schematic depiction of complex plastids with focus on membrane architecture and plastid preprotein import. Organisms evolved through the engulfment of either a red (**a**, Cryptophytes; **b**, Heterokontophytes and Haptophytes; **c**, Apicomplexa; **e**, peridinin-containing Dinoflagellates) or green alga (**d**, Chlorarachniophytes; and **e**, Euglenophytes) in an event called secondary endosymbiosis. These organisms harbor complex plastids with diverse membrane architecture and therefore different preprotein import pathways. (**a–d)** Plastids are surrounded by four membranes and the outermost membrane is either in continuum with the ER membrane (**a**, Cryptophytes; **b**, Heterokontophytes and Haptophytes) or not (**c,** Apicomplexa and **d** Chlorarachniophytes) (for a detailed description see text). In the first case (**a + b**) translocation over the first membrane is mediated by the Sec61 complex. If the outermost membrane is not connected with the ER (**c + d**), proteins are transported to the plastid through the endomembrane system—with or without involvement of the Golgi. (**e**) Euglenophytes and peridinin-containing dinoflagellates have plastids bounded by three membranes. Again, protein import is accomplished with a vesicle transport through the endomembrane system—in this case *via* ER and Golgi. *Nu* nucleus, *Nm* nucleomorph, *ER* endoplasmic reticulum, *Go* Golgi apparatus, *cER* chloroplast ER, *PPC* periplastidal compartment, *PPM* periplastidal membrane, *IMS* intermembrane space, *Sec61* ER Sec61-translocation complex, *SELMA* symbiont-specific ERAD-like translocation machinery, *TOC* translocon of the outer chloroplast membrane, *TIC* translocon of the inner chloroplast membrane. Uncertain processes are indicated by *question marks*

Despite different numbers of surrounding membranes and phylogenetic divergence, nucleus-encoded plastid preproteins share a basic targeting concept in such a way that preprotein transport into complex plastids generally starts cotranslationally at the ER membrane (Hempel et al. 2007; Bolte et al. 2009). In line with this targeting concept, these proteins are encoded with an N-terminal bipartite targeting signal (BTS) consisting of a signal peptide and a transit peptide-like sequence to accomplish proper preprotein import into the plastid.

Nevertheless, two different routes have to be distinguished as in heterokontophytes, haptophytes and cryptophytes the outermost plastid membrane is continuous with the ER membrane (Hempel et al. 2007; Bolte et al. 2009). Thus, cotranslational import into the ER is the mechanism to cross the outermost plastid membrane. Other scenarios apply in case of euglenophytes, peridinin-containing dinoflagellates, chlorarachniophytes, and perhaps in apicomplexa, in which the outermost membranes of the plastids are not connected with the ER. Nevertheless, the route of protein transport into the plastid starts cotranslationally at the ER membrane as well. As a consequence, vesicles have to act as transport shuttle between the secretory system and the outermost plastid membrane. Vesicle fusion with this membrane represents the first step in importing preproteins into the complex plastid. These vesicles might be guided from the ER directly to the first membrane in case of apicomplexa (DeRocher et al. 2000, 2005; Tonkin et al. 2006) (for a possible different view, see later), whereas the Golgi apparatus is traversed in euglenophytes and peridinin-containing dinoflagellates (Osafune et al. 1991; Sulli et al. 1999; Nassoury et al. 2003; Sláviková et al. 2005). Data providing information about protein transport into the complex plastids of chlorarachniophytes are hitherto limited; thus, it is still unclear if vesicles destined for the plastids originate at the Golgi or the ER membrane (Hirakawa et al. 2010, 2012) (Fig. 1).

In organisms with plastids surrounded by four membranes, the second outermost membrane is called periplastidal membrane (PPM) and separates the ER lumen from the so-called periplastidal compartment (PPC), the remnant cytoplasm of the secondary endosymbiont. For preprotein transport across this membrane, it was shown that an ERAD-like translocation machinery mediates this transport step in organisms of the red lineage with plastids surrounded by four membranes (Sommer et al. 2007; Hempel et al. 2009, 2010; Agrawal et al. 2009; Kalanon et al. 2009; Spork et al. 2009). A recent study failed to identify subunits of that ERAD-like translocation machinery in the similarly organized plastids of chlorarachniophytes (Hirakawa et al. 2012), and therefore it is of huge interest how these organisms manage preprotein transport across this membrane. In euglenophytes and peridinin-containing dinoflagellates, the respective membrane was probably lost during evolution (Schnepf and Deichgräber 1984; Patron et al. 2005; Lukeš et al. 2009).

The third and the fourth plastid membrane in the red lineage and in chlorarachniophytes are homologous to the outer and inner envelope membrane of chloroplasts. Both contain proteins homologous to components of the TOC and TIC translocon, respectively (Douglas et al. 2001; Gilson et al. 2006; van Dooren et al. 2008; Bullmann et al. 2010; Hirakawa et al. 2012). Accordingly, transport across the two inner membranes of these secondary plastids seems to be conserved with respect to primary plastids (Kessler and Schnell 2009; Bolte et al. 2009). In case of three membrane-bound plastids (euglenophytes and peridinin-containing dinoflagellates), less genomic data is available, and therefore candidates for TOC/TIC homologs have not been identified yet. But it seems feasible that these translocons of the former primary plastid are conserved in these organisms as well (Fig. 1). Whole genome sequences of these organisms will help to resolve this issue.

Open Questions on Preprotein Transport into Complex Plastids

Crossing the Outermost Membrane

In cryptophytes, haptophytes, and heterokontophytes, soluble proteins of the complex plastid are transported cotranslationally via Sec61 into the ER lumen and by this way also across the outermost membrane of the plastid. This scenario is completely different in those groups which have not included their plastids into the ER. Here, the first step—the translation into the rER—is not equivalent to the translocation across the outermost plastid membrane. However, in any case discrimination between secretory and plastid preproteins has to occur.

As mentioned, nucleus-encoded plastid preproteins of peridinin-containing dinoflagellates and euglenophytes traverse ER and Golgi on their route into the secondary plastid. Thus, one would expect a protein-sorting mechanism in the Golgi compartment, which might be similar to that known for differential sorting at the *trans*-Golgi (e.g., Bonifacino 2004). However, these sorting mechanisms are neither known for peridinin-containing dinoflagellates nor euglenophytes so far. If the signal peptide of the BTS is cleaved within the ER lumen as expected, the transit peptide-like part of the BTS can act as such a discrimination signal. But it is unknown if the transit peptide alone is sufficient to mediate the recruitment to vesicles and targeting towards the plastid.

Although phylogenetically not related, euglenophytes and peridinin-containing dinoflagellates share comparable additional features in their BTS (Patron et al. 2005; Durnford and Gray 2006). Several analyzed proteins contain an additional hydrophobic domain, C-terminal to their transit peptide-like sequence, which might act as a thylakoid-targeting domain and/or stop-transfer domain. As some of the investigated proteins are anchored in vesicle membranes via this hydrophobic domain in an N-in/C-out topology, i.e., the mature protein is exposed to the cytosol (Sulli et al. 1999; Nassoury et al. 2003), further targeting signals might be hidden in the mature part of the proteins.

For apicomplexa, two different models exist explaining targeting of preproteins to the four membrane-bound apicoplast. In the first model (model A, Tonkin et al. 2008a), the apicoplast is located within the ER, whereas in the second (model B) the complex plastid of the apicomplexa is in near contact, but physically independent from the ER of the host cell. In model A, apicoplast-directed preproteins have to be separated from secretory proteins and transported selectively across the second outermost apicoplast membrane like in cryptophytes, haptophytes, and heterokontophytes. In model B, the sorting occurs in the ER lumen as well, but additionally a vesicle flow from the ER to the outer apicoplast membrane is needed. This model is supported by the detection of vesicles transporting proteins for the apicoplast (Karnataki et al. 2007; DeRocher et al. 2008; van Dooren et al. 2008) (Fig. 1c). For both models, the transit peptide

might be an important signal for the discrimination of secretory proteins and apicoplast preproteins as there are no differences in the signal peptide composition of plastidal and secretory proteins known so far. And indeed, targeting experiments in apicomplexa resulted in the finding that the deletion of the transit peptide-like sequence of the BTS led to secretion of the respective proteins (e.g., Waller et al. 2000; DeRocher et al. 2000; Foth et al. 2003).

As protein turnover is expected to be high in the plastid, the vesicle flow to the plastid should be significant in organisms harboring secondary plastids whose outer membrane is not connected to the ER of the host cell. In order to prevent overloading of the outer membrane with lipids, a recycling mechanism, either by vesicles or lipid-transporting proteins, should be present.

Furthermore, the second outermost membrane has to be supplied with lipids in all organisms harboring secondary plastids, because these might not be provided by preprotein transporting vesicles. Mechanistic details still remain elusive.

Taken together, crossing the outermost membrane seems to be mechanistically more complex in chlorarachniophytes, euglenophytes, peridinin-containing dinoflagellates, and (in case of model B) in apicomplexa than in organisms, in which the outermost membrane is continuous with the ER of the host cell and preproteins cross the first membrane cotranslationally via Sec61.

Integration of Proteins into the Outermost Membrane

In the case of apicomplexa (in line with model B, see above), nucleus-encoded plastidal membrane proteins, which are supposed to get integrated into the outermost plastid membrane, might need specific targeting signals to distinguish them from membrane proteins designated for the secretory pathway. And indeed, proteins designated for the outer apicoplast membrane were recently described to lack a bipartite targeting signal (BTS), the general targeting signal for nucleus-encoded plastid/apicoplast preproteins [reviewed in Lim et al. (2009); Lim (2010)]. As a common targeting signal for those proteins, a signal anchor might be present, but this anchor would mediate the integration of the proteins into the ER membrane. Therefore, a second, still unknown targeting signal responsible for vesicle integration and targeting to the outermost membrane might exist. Accordingly, a similar targeting of proteins of the outermost membrane of plastids surrounded by three membranes seems to be necessary, but less is known about the protein composition of the outermost membrane in peridinin-containing dinoflagellates and euglenophytes.

These two groups face an additional problem at the outermost membrane resulting from the described architecture of their bipartite targeting signals. Preproteins, which appear to get anchored in the ER membrane (Sulli et al. 1999; Nassoury et al. 2003), and thereby in vesicle membranes, should be integrated into the outermost membrane after vesicle fusion. Predicted stromal proteins of these two protist groups often possess a membrane anchor as well, which would integrate

them in the outermost plastid membrane after vesicle fusion (Sulli et al. 1999; Patron et al. 2005; Sláviková et al. 2005; Durnford and Gray 2006). Thus, they have to be released from this membrane to be further transported into the stroma by an unknown mechanism. One concept might be that an unknown translocon, probably a member of the TOC family, in the second outermost membrane binds the exposed transit peptide and subsequently promotes the release and further transport of the preprotein. Another idea suggests that proteinaceous factors of the outer membrane attach to the preprotein and form a pore around it, which then enables the release of the protein (Sulli et al. 1999; van Dooren et al. 2001). However, these mechanisms would have to promote a probably complex membrane extraction of the whole mature protein, which is exposed to the cytoplasm after vesicle fusion with the outermost membrane. An anchorage of the protein in the membrane seems to increase the intricacy of the whole transport process and the discovery of the mechanistic solution of this problem would be quite interesting. Otherwise, we might have to rethink this model in so far that a much more simple mechanism might apply right at the beginning of the transport process.

The Role of the N-Terminal Targeting Signals

With the exceptions mentioned, nucleus-encoded preproteins targeted to secondary plastids harbor an N-terminal targeting signal, a bipartite targeting signal, BTS (Patron and Waller 2007). According to the current model of protein transport in cryptophytes and heterokontophytes, the signal peptide is cleaved off (by a yet unknown signal peptide peptidase) in the course of or after passing the outermost plastid membrane (Hempel et al. 2007; Gould et al. 2008; Bolte et al. 2009). Thus, as in the case of apicomplexa, these plastid preproteins can be sorted out from proteins of the secretory pathway in the ER lumen via the N-terminal transit peptide-like sequence. This implies the necessity for a receptor-like molecule in the space between the first and second outermost membrane with the capacity to direct the proteins to the translocon of the second outermost membrane. However, there is no clear evidence that the signal peptide is cleaved off in the ER lumen at all. Alternatively, the signal peptide might remain uncleaved, and the BTS would act as a whole in targeting to and translocation across the second membrane.

Recent investigations indicate that the overall charge of the transit peptide is critical for import into complex plastids regardless of the position of the charged amino acids (Foth et al. 2003; Tonkin et al. 2006, 2008b; Felsner et al. 2010). Surprisingly, one of these investigations shows furthermore that crossing the second as well as the two innermost membranes of diatoms depends on different ratios of positively charged amino acids of the transit peptide (Felsner et al. 2010).

Further important issues in discrimination of plastid preproteins and secretory proteins in the lumen between the first and second outermost membrane might be protein folding and posttranslational modifications. Proteins destined for the secretory pathway are folded after their cotranslational insertion into the ER via Sec61,

and many of them get glycosylated. However, we assume that folding has to be prevented for nucleus-encoded stromal preproteins of heterokontophytes, haptophytes, and apicomplexa, in order to keep proteins in a transport-competent conformation for crossing the third outermost membrane via conserved Toc75 proteins (Bullmann et al. 2010). If so, one would expect the same mechanism and factors in cryptophytes, but even newly available genome data for the cryptophyte *Guillardia theta* could not help to identify a Toc75 homolog so far (Curtis et al. 2012). Nevertheless, in cryptophytes, a TOC system seems to be the most reasonable candidate for translocation over the third plastid membrane.

Transport of Membrane Proteins Destined for the Second, Third, or Fourth Membrane

Several nucleus-encoded membrane proteins are inserted into the membranes surrounding the secondary symbiont as exemplarily shown in the diatom *Phaeodactylum tricornutum* for plastid proteins located in the second and third outermost membranes (Hempel et al. 2009, 2010; Bullmann et al. 2010; Moog et al. 2011; Stork et al. 2012). With the exception of the beta-barrel protein Toc75, the proposed translocon in the third outermost membrane (Bullmann et al. 2010), the other known symbiont membrane proteins are predicted to contain membrane-spanning α-helices (Hempel et al. 2009, 2010; Stork et al. 2012). However, α-helical membrane-spanning domains should have the affinity to integrate into the outermost membrane while being imported via Sec61. This might be prevented in case of integral membrane proteins destined for the second, third, or fourth membrane of the complex plastid to ensure integration into the correct downstream membrane. It is reasonable to expect specific signals for such a complicated targeting, even though these are not known at the moment. Compared to soluble preproteins, alternative mechanisms like vesicle-mediated transport might apply as well.

PPC Versus Stroma Targeting in Plastids Surrounded by Four Membranes

In general, soluble nucleus-encoded proteins for the symbiont can be classified into (1) proteins crossing two membranes to reach the PPC, (2) proteins crossing three membranes to enter the intermembrane space between the third and fourth membranes (IMS), and (3) stroma and thylakoid directed proteins, which have to pass either four or five membranes (Fig. 1).

For crossing the second outermost membrane, a preprotein translocation machinery was described (Sommer et al. 2007; Hempel et al. 2009, 2010). This

machinery was identified in all secondarily evolved organisms with a symbiont of red algal origin except for peridinin-containing dinoflagellates (Sommer et al. 2007; Spork et al. 2009; Kalanon et al. 2009; Agrawal et al. 2009; Felsner et al. 2011; Stork et al. 2012). It is proposed to be a modified ERAD translocation system, which was shown—in the case of the experimentally most suitable apicomplexan parasite *Toxoplasma gondii*—to be involved in preprotein transport into the complex plastid (Agrawal et al. 2009). Such a protein translocation system should be necessary for all plastidal preproteins crossing the second outermost membrane. Surprisingly, in chlorarachniophytes, no such symbiont-specific ERAD system could be identified (Hirakawa et al. 2012). In future, it will be very interesting to find out how preprotein transport across the respective membrane is managed in these organisms. The genome project of the chlorarachniophyte *Bigelowiella natans* might give further insights on that issue (Curtis et al. 2012).

For further transport across the PPC, two scenarios are possible (Cavalier-Smith 1999): either plastid preproteins cross the second membrane completely and are then directed to a proteinaceous translocon in the third membrane, or preproteins pass the PPC in one single step via two connected translocons of membranes two and three. The latter might be mechanistically similar to the coupled transport known from the contact sites of mitochondrial protein import via TOM and TIM (e.g., Reichert and Neupert 2002). The recent identification of the putative translocon in the third outermost membrane, a Toc75 homolog (Bullmann et al. 2010), does not necessarily favor one possibility. However, at least in cryptophytes and chlorarachniophytes, which still harbor a protein synthesis machinery inside the PPC expressing some plastid proteins encoded on the vestigial genome of the eukaryotic symbiont (Douglas et al. 2001; Gilson et al. 2006; Lane et al. 2007; Moore et al. 2012), a protein translocation machinery must be present for transport across membrane three and four, which is independent from protein transport across membrane two. Crossing these two membranes is most probably also facilitated by a TOC/TIC system since TIC and, in case of chlorarachniophytes, TOC components have been identified in these two lineages (Douglas et al. 2001; Gilson et al. 2006; Hirakawa et al. 2012).

A comparable situation is given for PPC-resident proteins encoded in the host nucleus. In heterokontophytes and cryptophytes, those proteins are discriminated from plastid preproteins via the +1 amino acid position of the transit peptide-like sequence (Gould et al. 2006a, b; Gruber et al. 2007). In the case of plastid preproteins, this position is an aromatic amino acid or in some cases a leucine, whereas PPC proteins do have other amino acids at this position in heterokontophytes and cryptophytes. Many results obtained by the investigation of the symbiont-specific ERAD-like translocation machinery (SELMA) (Hempel et al. 2009, 2010; Stork et al. 2012) can be explained by both routes of entry, but nevertheless independent machineries in the second and third membranes, i.e., no coupled transport between the ERAD-like machinery and Toc75, are more likely.

Some years ago, another alternative model was suggested for proteins with a destination beyond the PPC, i.e., for IMS, stromal or thylakoid proteins (Gibbs 1979; Kilian and Kroth 2005). Here, a vesicle flow between the second and third

outermost membranes was postulated. One reason for this hypothesis was the observation of membranous structures within the PPC in electron microscopic studies (Gibbs 1979). However, a recent genomic screen accomplished in our lab failed to identify PPC located candidate proteins known to be essential for vesicle budding, transport, and fusion (Moog et al. 2011). Thus, the observed structures might have other functions than protein transport.

Although chlorarachniophytes have a symbiont of green algal origin, protein transport might be in some aspects similar to import into plastids of organisms with a red algal-derived symbiont. This can be inferred from the bipartite targeting signal, which is located at the N-terminus of chlorarachniophyte nucleus-encoded plastid proteins as well (Rogers et al. 2004; Hirakawa et al. 2009), and the presence of a Toc75 homolog, encoded by the minimized genome of the eukaryotic symbiont in the PPC of chlorarachniophytes (Gilson et al. 2006; Hirakawa et al. 2012). However, the transit peptide-like sequences of plastid-targeted preproteins of chlorarachniophytes are not compatible with those of apicomplexa and cryptophytes, which indicates important differences in intracellular targeting (Hirakawa et al. 2009). In addition, at least some nucleus-encoded proteins directed into the PPC of chlorarachniophytes harbor an additional C-terminal signal, which might act as a PPC-retention signal (Hirakawa et al. 2010). A further hint for unconventional targeting in chlorarachniophytes represents the targeting of the RuBisCo small subunit protein, for which it was shown that not only the N-terminal BTS but also a signal located within the mature protein is necessary for correct delivery of the protein into the stroma (Hirakawa and Ishida 2010).

Recycling the Old for Generating the New

Recent investigations on protein transport into complex plastids revealed that preexisting translocation machineries were “recycled” and modified during evolution to serve new functions in preprotein transport (Bolte et al. 2011).

In the case of cryptophytes, haptophytes, heterokontophytes, and apicomplexa, the ERAD-transport machinery of the red algal endosymbiont is thought to be relocated to the second outermost plastid membrane and established as a preprotein translocator (Sommer et al. 2007; Hempel et al. 2009, 2010; Spork et al. 2009; Kalanon et al. 2009; Agrawal et al. 2009; Felsner et al. 2011; Stork et al. 2012). Components for the translocation machineries of the two innermost membranes have been identified in haptophytes, heterokontophytes, apicomplexa, and chlorarachniophytes so far (Rogers et al. 2007; van Dooren et al. 2008; Kalanon et al. 2009; Bullmann et al. 2010; Hirakawa et al. 2012) (Fig. 1). These seem to be conserved to the protein import machineries of primary plastids, namely TOC and TIC, which are quite well investigated in chloroplasts of land plants (Inaba and Schnell 2008). The Toc75 homolog in the third plastid membrane of heterokontophytes, the protein ptOmp85, was recently characterized concerning targeting mechanisms and electrophysiological properties (Bullmann et al. 2010). PtOmp85

was shown to recapitulate the complex targeting mechanism of Toc75 proteins from land plants by first traversing all four membranes of the complex plastid into the stroma, and thereafter being retrotranslocated and integrated into the third outermost membrane.

Another indication for the conservation of protein import across the two innermost membranes is a conserved aromatic amino acid found at the very N-terminus of the transit peptide of stromal proteins of heterokontophytes and cryptophytes (Kilian and Kroth 2005; Gould et al. 2006a, b; Gruber et al. 2007; Patron and Waller 2007). In the PPC, imported preproteins are discriminated by the first amino acid of their transit peptides: preproteins with an aromatic amino acid [predominantly a phenylalanine (Phe) or a leucine (Leu)] at this position are further transported across the two innermost membranes, whereas preproteins without such a characteristic aromatic amino acid are retained in the PPC. A similar phenylalanine-dependency was observed for the protein import into the primary plastids of glaucophytes, and it was shown that this dependency is most likely due to a direct interaction of the phenylalanine and the Toc75 translocon in the outer plastid membrane (Wunder et al. 2007). In contrast, a comparable phenylalanine-dependency is not conserved in chlorophytes, and hence, the conspicuous amino acid is missing in transit peptides of chlorophytes as well as in chlorarachniophytes and euglenophytes (Patron and Waller 2007). Thus, although preprotein transport across the two innermost membranes of all investigated complex plastids of red and green algal origin might use a TOC/TIC system, differences in cargo reception and/or in transport mechanisms might be present depending on the nature of the respective endosymbiont.

Vesicle transport for plastid preprotein import is another feature, which is not unique for complex plastids. Also in land plants, it has been shown that the transport of several nucleus-encoded proteins into the primary plastid depends on vesicle transport involving the endomembrane system (Villarejo et al. 2005, Nanjo et al. 2006, Kitajima et al. 2009). Even though detailed mechanisms are not well understood in both cases so far, it seems quite convincing that vesicle transport is an elegant way to deal with protein transport issues.

A further class of plastid proteins that might share similar targeting mechanisms in primary and secondary plastids are nucleus-encoded proteins of the thylakoids. Experimental data concerning thylakoid import in complex plastids are rare, but the integration of a diatom Fcp protein into thylakoid membranes, as well as the transport of a cryptophyte phycoerythrin subunit across the thylakoid membrane via the TAT pathway was shown in homologous and heterologous in vitro systems (Lang and Kroth 2001; Gould et al. 2007). In the case of complex plastids surrounded by three membranes, thylakoid import of the protein PsbO of a peridinin-containing dinoflagellate was studied by import assays using pea chloroplasts (Chaal and Green 2005). All these experiments led to the conclusion that integration into and transport across thylakoid membranes follow similar principles as in land plants.

Conclusions

In many secondarily evolved organisms, preprotein import into the complex plastid has turned out to be mediated by a combination of preexisting "recycled" mechanisms. However, even though the translocation machineries for transport of nucleus-encoded plastid proteins seem to be identified for organisms with red algal derived plastids surrounded by four membranes, many functional aspects are still not known. On the other side, basic knowledge on preprotein import into plastids of green algal origin and peridinin-containing dinoflagellates is still lacking in many aspects, as for these organisms less experimental strategies are available. New insights into preprotein targeting into plastids surrounded by three membranes will help additionally in defining the origin of these membranes. Characterization of protein transport strategies across the outermost and second outermost plastid membrane of chlorarachniophytes will indicate how these organisms facilitate the transport of preproteins across these membranes. At least crossing the second outermost membrane seems to be different from the red lineage as no SELMA equivalent could be identified (Hirakawa et al. 2012). Resolving the mechanisms of this transport step might uncover new evolutionary strategies dealing with plastid protein import.

Acknowledgments We are supported by the Deutsche Forschungsgemeinschaft within SFB 593, SFB-TR1, and GRK 1216, and the LOEWE program of the State of Hessen.

References

Agrawal S, van Dooren GG, Beatty WL, Striepen B (2009) Genetic evidence that an endosymbiont-derived endoplasmic reticulum-associated protein degradation (ERAD) system functions in import of apicoplast proteins. J Biol Chem 284:33683–33691

Archibald JM (2009) The puzzle of plastid evolution. Curr Biol 19:81–88

Bionda T, Tillmann B, Simm S, Beilstein K, Ruprecht M, Schleiff E (2010) Chloroplast import signals: the length requirement for translocation in vitro and in vivo. J Mol Biol 402:510–523

Bolte K, Bullmann L, Hempel F, Bozarth A, Zauner S, Maier UG (2009) Protein targeting into secondary plastids. J Eukaryot Microbiol 56:9–15

Bolte K, Gruenheit N, Felsner G, Sommer MS, Maier UG, Hempel F (2011) Making new out of old: recycling and modification of an ancient protein translocation system during eukaryotic evolution. Bioessays 33(5):368–376

Bonifacino JS (2004) The GGA proteins: adaptors on the move. Nat Rev Mol Cell Biol 5:23–32

Bullmann L, Haarmann R, Mirus O, Bredemeier R, Hempel F, Maier UG, Schleiff E (2010) Filling the gap, evolutionarily conserved Omp85 in plastids of chromalveolates. J Biol Chem 285:6848–6856

Cavalier-Smith T (1999) Principles of protein and lipid targeting in secondary symbiogenesis: euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree. J Eukaryot Microbiol 46:347–366

Cavalier-Smith T (2003) Genomic reduction and evolution of novel genetic membranes and protein-targeting machinery in eukaryote-eukaryote chimaeras (meta-algae). Philos Trans R Soc Lond B Biol Sci 358:109–133

Chaal BK, Green BR (2005) Protein import pathways in 'complex' chloroplasts derived from secondary endosymbiosis involving a red algal ancestor. Plant Mol Biol 57:333–342

Curtis BA, Tanifuji G, Burki F, Gruber A, Irimia M, Maruyama S, Arias MC, Ball SG, Gile GH, Hirakawa Y, Hopkins JF, Kuo A, Rensing SA, Schmutz J, Symeonidi A, Elias M, Eveleigh RJM, Herman EK, Klute MJ, Nakayama T, Obornik M, Reyes-Prieto A, Armbrust EV, Aves SJ, Beiko RG, Coutinho P, Dacks JB, Durnford DG, Fast NM, Green BR, Grisdale CJ, Hempel F, Henrissat B, Hoppner MP, Ishida K-I, Kim E, Koreny L, Kroth PG, Liu Y, Malik S-B, Maier UG, McRose D, Mock T, Neilson JAD, Onodera NT, Poole AM, Pritham EJ, Richards TA, Rocap G, Roy SW, Sarai C, Schaack S, Shirato S, Slamovits CH, Spencer DF, Suzuki S, Worden AZ, Zauner S, Barry K, Bell C, Bharti AK, Crow JA, Grimwood J, Kramer R, Lindquist E, Lucas S, Salamov A, McFadden GI, Lane CE, Keeling PJ, Gray MW, Grigoriev IV, Archibald JM (2012) Algal genomes reveal evolutionary mosaicism and the fate of nucleomorphs. Nature 492:59–65

DeRocher A, Hagen CB, Froehlich JE, Feagin JE, Parsons M (2000) Analysis of targeting sequences demonstrates that trafficking to the *Toxoplasma gondii* plastid branches off the secretory system. J Cell Sci 113:3969–3977

DeRocher A, Gilbert B, Feagin JE, Parsons M (2005) Dissection of brefeldin A-sensitive and -insensitive steps in apicoplast protein targeting. J Cell Sci 118:565–574

DeRocher AE, Coppens I, Karnataki A, Gilbert LA, Rome ME, Feagin JE, Bradley PJ, Parsons M (2008) Thioredoxin family protein of the apicoplast periphery identifies abundant candidate transport vesicles in *Toxoplasma gondii*. Eukaryot Cell 7:1518–1529

Douglas S, Zauner S, Fraunholz M, Beaton M, Penny S, Deng LT, Wu X, Reith M, Cavalier-Smith T, Maier UG (2001) The highly reduced genome of an enslaved algal nucleus. Nature 410:1091–1096

Durnford DG, Gray MW (2006) Analysis of *Euglena gracilis* plastid- targeted proteins reveals different classes of transit sequences. Eukaryot Cell 5:2079–2091

Felsner G, Sommer MS, Maier UG (2010) The physical and functional borders of transit peptide-like sequences in secondary endosymbionts. BMC Plant Biol 10:223

Felsner G, Sommer MS, Gruenheit N, Hempel F, Moog D, Zauner S, Martin W, Maier UG (2011) ERAD components in organisms with complex red plastids suggest recruitment of a preexisting protein transport pathway for the periplastid membrane. Genome Biol Evol 3:140–150

Foth BJ, Ralph SA, Tonkin CJ, Struck NS, Fraunholz M, Roos DS, Cowman AF, McFadden GI (2003) Dissecting apicoplast targeting in the malaria parasite Plasmodium falciparum. Science 299:705–708

Gibbs SP (1979) The route of entry of cytoplasmically synthesized proteins into chloroplasts of algae possessing chloroplast ER. J Cell Sci 35:253–266

Gilson PR, Su V, Slamovits CH, Reith ME, Keeling PJ, McFadden GI (2006) Complete nucleotide sequence of the chlorarachniophyte nucleomorph: nature's smallest nucleus. Proc Natl Acad Sci U S A 103:9566–9571

Gould SB, Sommer MS, Hadfi K, Zauner S, Maier UG (2006a) Protein targeting into the complex plastid of cryptophytes. J Mol Evol 62:674–681

Gould SB, Sommer MS, Kroth PG, Gile GH, Keeling PJ, Maier UG (2006b) Nucleus-to-nucleus gene transfer and protein retargeting into a remnant cytoplasm of cryptophytes and diatoms. Mol Biol Evol 23:2413–2422

Gould SB, Fan E, Hempel F, Maier UG, Klösgen RB (2007) Translocation of a phycoerythrin alpha subunit across five biological membranes. J Biol Chem 282:30295–30302

Gould SB, Waller RF, McFadden GI (2008) Plastid evolution. Annu Rev Plant Biol 59:491–517

Gruber A, Vugrinec S, Hempel F, Gould SB, Maier UG, Kroth PG (2007) Protein targeting into complex diatom plastids: functional characterisation of a specific targeting motif. Plant Mol Biol 64:519–530

Hempel F, Bozarth A, Sommer MS, Zauner S, Przyborski JM, Maier UG (2007) Transport of nuclear-encoded proteins into secondarily evolved plastids. Biol Chem 388:899–906

Hempel F, Bullmann L, Lau J, Zauner S, Maier UG (2009) ERAD derived preprotein transport across the second outermost plastid membrane of diatoms. Mol Biol Evol 26:1781–1790
Hempel F, Felsner G, Maier UG (2010) New mechanistic insights into pre-protein transport across the second outermost plastid membrane of diatoms. Mol Microbiol 76:793–801
Hirakawa Y, Ishida K (2010) Internal plastid-targeting signal found in a RuBisCO small subunit protein of a chlorarachniophyte alga. Plant J 64:402–410
Hirakawa Y, Nagamune K, Ishida K (2009) Protein targeting into secondary plastids of chlorarachniophytes. Proc Natl Acad Sci U S A 106:12820–12825
Hirakawa Y, Gile GH, Ota S, Keeling PJ, Ishida K (2010) Characterization of periplastidal compartment-targeting signals in chlorarachniophytes. Mol Biol Evol 27:1538–1545
Hirakawa Y, Burki F, Keeling PJ (2012) Genome-based reconstruction of the protein import maschinery in the secondary plastid of a chlorarachniophyte alga. Eukaryot Cell 11:324–333
Inaba T, Schnell DJ (2008) Protein trafficking to plastids: one theme, many variations. Biochem J 413:15–28
Janouškovec J, Hórak A, Oborník M, Lukeš J, Keeling PJ (2011) A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids. Proc Natl Acad Sci USA 107:10949–10954
Kalanon M, Tonkin CJ, McFadden GI (2009) Characterization of two putative protein translocation components in the apicoplast of *Plasmodium falciparum*. Eukaryot Cell 8:1146–1154
Karnataki A, Derocher A, Coppens I, Nash C, Feagin JE, Parsons M (2007) Cell cycle-regulated vesicular trafficking of *Toxoplasma* APT1, a protein localized to multiple apicoplast membranes. Mol Microbiol 63:1653–1668
Keeling PJ (2009) Chromalveolates and the evolution of plastids by secondary endosymbiosis. J Eukaryot Microbiol 56:1–8
Kessler F, Schnell D (2009) Chloroplast biogenesis: diversity and regulation of the protein import apparatus. Curr Opin Cell Biol 21:494–500
Kilian O, Kroth PG (2005) Identification and characterization of a new conserved motif within the presequence of proteins targeted into complex diatom plastids. Plant J 41:175–183
Kitajima A, Asatsuma S, Okada H, Hamada Y, Kaneko K, Nanjo Y, Kawagoe Y, Toyooka K, Matsuoka K, Takeuchi M, Nakano A, Mitsui T (2009) The rice alpha-amylase glycoprotein is targeted from the Golgi apparatus through the secretory pathway to the plastids. Plant Cell 21:2844–2858
Lane CE, van den Heuvel K, Kozera C, Curtis BA, Parsons BJ, Bowman S, Archibald JM (2007) Nucleomorph genome of *Hemiselmis andersenii* reveals complete intron loss and compaction as a driver of protein structure and function. Proc Natl Acad Sci U S A 104:19908–19913
Lang M, Kroth PG (2001) Diatom fucoxanthin chlorophyll a/c-binding protein (FCP) and land plant light-harvesting proteins use a similar pathway for thylakoid membrane insertion. J Biol Chem 276:7985–7991
Lim L, McFadden GI (2010) The evolution, metabolism and functions of the apicoplast. Philos Trans R Soc Lond B Biol Sci 365:749–763
Lim L, Kalanon M, McFadden GI (2009) New proteins in the apicoplast membranes: time to rethink apicoplast protein targeting. Trends Parasitol 25:197–200
Lukeš J, Leander BS, Keeling PJ (2009) Cascades of convergent evolution: the corresponding evolutionary histories of euglenozoans and dinoflagellates. Proc Natl Acad Sci USA 106:9963–9970
Maier UG, Douglas SE, Cavalier-Smith T (2000) The nucleomorph genomes of cryptophytes and chlorarachniophytes. Protist 151:103–109
Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, Leister D, Stoebe B, Hasegawa M, Penny D (2002) Evolutionary analysis of Arabidopsis, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proc Natl Acad Sci U S A 99:12246–12251
Moog M, Stork S, Zauner S, Maier UG (2011) *In silico* and *in vivo* investigations of the proteins of a minimized eukaryotic cytoplasm. Genome Biol Evol 3:375–382

Moore CE, Curtis B, Mills T, Tanifuji G, Archibald JM (2012) Nucleomorph genome sequence of the cryptophyte alga *Chroomonas mesostigmatica* CCMP1168 reveals lineage-specific gene loss and genome complexity. Genome Biol Evol 4:1162–1175

Nassoury N, Cappadocia M, Morse D (2003) Plastid ultrastructure defines the protein import pathway in dinoflagellates. J Cell Sci 116:2867–2874

Nanjo Y, Oka H, Ikarashi N, Kaneko K, Kitajima A, Mitsui T, Muñoz FJ, Rodríguez-López M, Baroja-Fernández E, Pozueta-Romero J (2006) Rice plastidial N-glycosylated nucleotide pyrophosphatase/phosphodiesterase is transported from the ER-golgi to the chloroplast through the secretory pathway. Plant Cell 18:2582–2592

Osafune T, Schiff JA, Hase E (1991) Stage-dependent localization of LHCP II apoprotein in the Golgi of synchronized cells of *Euglena gracilis* by immunogold electron microscopy. Exp Cell Res 193:320–330

Patron NJ, Waller RF (2007) Transit peptide diversity and divergence: a global analysis of plastid targeting signals. Bioessays 29:1048–1058

Patron NJ, Waller RF, Archibald JM, Keeling PJ (2005) Complex protein targeting to dinoflagellate plastids. J Mol Biol 348:1015–1024

Reichert AS, Neupert W (2002) Contact sites between the outer and inner membrane of mitochondria-role in protein transport. Biochim Biophys Acta 1592:41–49

Rogers MB, Archibald JM, Field MA, Striepen B, Keeling PJ (2004) Plastid-targeting peptides from the chlorarachniophyte *Bigelowiella natans*. J Eukaryot Microbiol 51:529–535

Rogers MB, Gilson PR, Su V, McFadden GI, Keeling PJ (2007) The complete chloroplast genome of the chlorarachniophyte *Bigelowiella natans*: evidence for independent origins of chlorarachniophyte and euglenid secondary endosymbionts. Mol Biol Evol 24:54–62

Schnepf E, Deichgräber G (1984) "Myzocytosis" a kind of endocytosis with implications to compartmentalization in endosymbiosis. Naturwissenschaften 71:218–219

Sláviková S, Vacula R, Fang Z, Ehara T, Osafune T, Schwartzbach SD (2005) Homologous and heterologous reconstitution of Golgi to chloroplast transport and protein import into the complex chloroplasts of *Euglena*. J Cell Sci 118:1651–1661

Soll J, Schleiff E (2004) Protein import into chloroplasts. Nat Rev Mol Cell Biol 5:198–208

Sommer MS, Gould SB, Lehmann P, Gruber A, Przyborski JM, Maier UG (2007) Der1-mediated pre-protein import into the periplastid compartment of chromalveolates? Mol Biol Evol 24:918–928

Spork S, Hiss JA, Mandel M, Sommer M, Kooij TWA, Chu T, Schneider G, Maier UG, Przyborski JM (2009) An unusual ERAD-like complex is targeted to the apicoplast of *Plasmodium falciparum*. Eukaryot Cell 8:1134–1145

Stork S, Moog D, Przyborski JM, Wilhelmi I, Zauner S, Maier UG (2012) Distribution of the SELMA translocon in secondary plastids of red algal origin and predicted uncoupling of ubiquitin-dependent translocation from degradation. Eukaryot Cell. doi:10.1128/EC.00183-12

Sulli C, Fang ZH, Muchhal U, Schwartzbach SD (1999) Topology of *Euglena* chloroplast protein precursors within endoplasmic reticulum to golgi to chloroplast transport vesicles. J Biol Chem 274:457–463

Tonkin CJ, Struck NS, Mullin KA, Stimmler LM, McFadden GI (2006) Evidence for Golgi-independent transport from the early secretory pathway to the plastid in malaria parasites. Mol Microbiol 61:614–630

Tonkin CJ, Kalanon M, McFadden GI (2008a) Protein targeting to the malaria parasite plastid. Traffic 9:166–175

Tonkin CJ, Roos DS, McFadden GI (2008b) N-terminal positively charged amino acids, but not their exact position, are important for apicoplast transit peptide fidelity in *Toxoplasma gondii*. Mol Biochem Parasitol 150:192–200

van Dooren GG, Schwartzbach SD, Osafune T, McFadden GI (2001) Translocation of proteins across the multiple membranes of complex plastids. Biochim Biophys Acta 1541:34–53

van Dooren GG, Tomova C, Agrawal S, Humbel BM, Striepen B (2008) *Toxoplasma gondii* Tic20 is essential for apicoplast protein import. Proc Natl Acad Sci U S A 105:13574–13579

Villarejo A, Burén S, Larsson S, Déjardin A, Monné M, Rudhe C, Karlsson J, Jansson S, Lerouge P, Rolland N, von Heijne G, Grebe M, Bako L, Samuelsson G (2005) Evidence for a protein transported through the secretory pathway en route to the higher plant chloroplast. Nat Cell Biol 7:1224–1231

Waller RF, Reed MB, Cowman AF, McFadden GI (2000) Protein trafficking to the plastid of *Plasmodium falciparum* is via the secretory pathway. EMBO J 19:1794–1802

Wunder T, Martin R, Löffelhardt W, Schleiff E, Steiner JM (2007) The invariant phenylalanine of precursor proteins discloses the importance of Omp85 for protein translocation into cyanelles. BMC Evol Biol 7:236

Tertiary Plastid Endosymbioses in Dinoflagellates

Przemysław Gagat, Andrzej Bodył, Paweł Mackiewicz, and John W. Stiller

Abstract Dinoflagellates are a peculiar group of protists with a surprising and varied history of plastid acquisition. They employ a variety of trophic strategies including photoautotrophy, heterotrophy, and mixotrophy, with multiple modes of food ingestion identified. This collection of features apparently preadapted dinoflagellates for acquisition of a bewildering array of photosynthetic bodies ranging from "stolen" plastids (or kleptoplastids) through permanent endosymbionts to true plastids, acquired in various primary, secondary, and tertiary endosymbioses. In this chapter, we focus on tertiary plastid endosymbioses (that is, uptake of an alga with a complex, secondary plastid), and especially on three that show distinct levels of host–endosymbiont integration. These endosymbiotic consortia are represented by (1) cryptophyte-derived kleptoplastids in *Dinophysis* species, (2) diatom endosymbionts in genera known as "dinotoms" (e.g., *Kryptoperidinium* and *Durinskia*), and (3) haptophyte-derived plastids in *Karenia*, *Karlodinium*, and *Takayama*. We discuss details of the structures, evolutionary origins, and processes involved in these varied endosymbioses, including feeding mechanisms, endosymbiotic gene transfer, and how nucleus-encoded proteins are targeted to each of these photosynthetic entities. Available data support previous predictions that all these photosynthetic bodies evolved via replacements of the peridinin plastid found in most photosynthetic dinoflagellates.

P. Gagat • P. Mackiewicz (✉)
Department of Genomics, Faculty of Biotechnology, University of Wrocław,
ul. Przybyszewskiego 63/77, 51-148 Wrocław, Poland
e-mail: pamac@smorfland.uni.wroc.pl

A. Bodył
Laboratory of Evolutionary Protistology, Department of Evolutionary Biology and Ecology,
University of Wrocław, ul. Przybyszewskiego 63/77, 51-148 Wrocław, Poland

J.W. Stiller
Department of Biology, East Carolina University, Howell Science, N108, Greenville,
NC 27858, USA

W. Löffelhardt (ed.), *Endosymbiosis*, DOI 10.1007/978-3-7091-1303-5_13,
© Springer-Verlag Wien 2014

Contents

Dinoflagellates: A Peculiar Protist Group with Experience in Plastid Acquisition 235
General Characteristics 235
"Core" Dinoflagellates: Their Characteristic Cellular and Molecular Features 235
Phylogeny of "Core" Dinoflagellates 238
Trophic Strategies in "Core" Dinoflagellates 239
Mixotrophy: A Trophic Strategy Widespread in Dinoflagellates 239
Feeding Mechanisms in Dinoflagellates 239
What Are the Differences Between Endosymbionts and Cell Organelles? 241
Main Kinds of Photosynthetic Bodies in "Core" Dinoflagellates: Peridinin Plastid, Unusual Plastids, and Kleptoplastids 241
Primary, Secondary, and Tertiary Plastid Endosymbioses in Dinoflagellates 241
Peridinin-Containing Plastid: The Most Typical Dinoflagellate Plastid 243
Unusual Dinoflagellate Plastids and Their Evolutionary Pathways 246
Kleptoplastids: A Peculiar Kind of Photosynthetic Bodies in Dinoflagellates 247
Cryptophyte-Derived Photosynthetic Bodies of *Dinophysis* Species: True Plastids or Kleptoplastids? 248
General Characteristic of *Dinophysis* 248
Evidence for Cryptophyte Origin of *Dinophysis* Chromatophores 250
Indirect *Myrionecta rubra*-Mediated Route of Cryptophyte Plastids into *Dinophysis* Cells 250
Nature of the Cryptophyte-Derived Chromatophores in *Dinophysis* Species: True Plastids or Kleptoplastids? 251
Protein Import into *Dinophysis* Kleptoplastids and the Origin of Their Envelope Membranes 254
The Reasons Behind *Dinophysis* Kleptoplastid Longevity 258
Nature of Plastid Replacement in *Dinophysis* Species 258
Dinoflagellates Containing Diatom-Derived Photosynthetic Bodies: Permanent Endosymbionts or True Cell Organelles? 259
General Characteristics of Dinotom Dinoflagellates 259
Ultrastructure and Origin of the Diatom-Derived Endosymbionts 260
Complex Evolutionary History of Dinotom Endosymbionts 262
Genomics of Diatom Endosymbiont Plastids 264
Putative Protein Import into Dinotom Endosymbionts 266
Peculiar Eyespot, a Remnant of the Peridinin Plastid? 268
True Plastids of Haptophyte Origin in Fucoxanthin Dinoflagellates 269
General Characteristics of *Kareniaceae* Dinoflagellates 269
Evidence for Haptophyte Origin of *Karenia* and *Karlodinium* Plastids 269
Karlodinium Plastid Genome and Its Transformation 271
How Many Membranes Surround Fucoxanthin Plastids: Two, Three, or Four? 272
Mode of Acquisition, Protein Targeting Mechanisms, and Membrane Homology of Fucoxanthin Plastids 273
Plastid Replacement in Fucoxanthin Dinoflagellates 274
Perspectives: Broader Significance of Tertiary Endosymbiosis 277
References 279

Dinoflagellates: A Peculiar Protist Group with Experience in Plastid Acquisition

General Characteristics

Dinoflagellates are predominantly unicellular protists, although sometimes filamentous or coenocytic, which live mainly in marine environments (Hackett et al. 2004a; Taylor 2004; Delwiche 2007; Moestrup and Daugbjerg 2007). They play important ecological roles as primary producers with a surprising diversity of photosynthetic bodies including plastids, endosymbionts, and kleptoplastids (for reviews, see Schnepf and Elbrächter 1999; Stoebe and Maier 2002; Nowack and Melkonian 2010). Dinoflagellates also are found as parasites, micrograzers, and as endosymbionts themselves (Gómez 2012a). For example, the genus *Symbiodinium* forms especially important symbiotic associations with reef-building corals (Baker 2003). Moreover, dinoflagellates are responsible for many harmful algal blooms, producing toxins with negative impacts on both wild and cultured animal populations, and which have serious economic consequences for fish and shellfish aquaculture, and for human health (Wang 2008; Brand et al. 2012; Reguera et al. 2012). To date, more than 2,000 extant dinoflagellate species have been described, along with a similar number of extinct taxa (Taylor et al. 2007).

"Core" Dinoflagellates: Their Characteristic Cellular and Molecular Features

The Dinoflagellata constitute a large and diverse protistan group (Fig. 1). They cluster with parasitic apicomplexans and, more distantly, with free-living ciliates (Leander and Keeling 2003; Bachvaroff et al. 2011). Apicomplexa contain a nonphotosynthetic plastid termed the apicoplast (for a review, see McFadden 2011), whereas ciliates are plastid-less protists. All these lineages, along with other protists, such as the genus *Perkinsus* (Saldarriaga et al. 2003) and chromerids (Moore et al. 2008; see also Linares et al. 2013), possess characteristic flattened vesicles known as alveoli that occur beneath their plasmalemma. Consequently, the whole superassembly is known as Alveolata (Cavalier-Smith 1993a).

The most basal branches of the dinoflagellate tree include the phagotrophic genus *Oxyrrhis* as well as parasitic lineages such as ellobiopsids and Marine Alveolate Groups (MAG) I and II, the latter corresponding to the Syndiniales (Fig. 1) (Moreira and López-Garcia 2002; Groisillier et al. 2006; Guillou et al. 2008; Gómez et al. 2009; Lowe et al. 2011). All later diverging species are known as "core" dinoflagellates (Fig. 1) (Gómez 2012b). The cells of "core" dinoflagellates have a characteristic transverse groove, or cingulum, which divides them into two parts, an upper epicone and a lower hypocone (Fig. 2). They swim

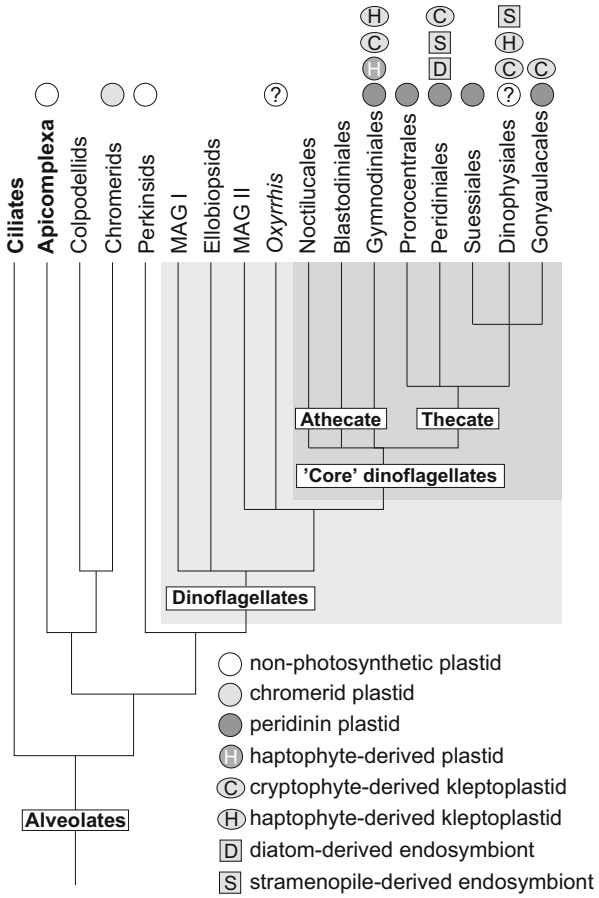


Fig. 1 Phylogeny of dinoflagellates and their relationships with other protist groups. Dinoflagellates belong to the supergroup Alveolata, named for the presence of flattened vesicles (or alveoli) that occur beneath the plasmalemma. The Alveolata also includes ciliates (characterized by nuclear dualism), parasitic apicomplexans (e.g., *Plasmodium falciparum*, the cause of malaria), colpodellids (predatory protists), photosynthetic chromerids (coral symbionts), and perkinsids (e.g., *Perkinsus marinus*, an oyster parasite). Dinoflagellata is recovered as the sister group to perkinsids in phylogenetic analyses. The mostly parasitic Marine Alveolate Group I (MAG I), MAG II (or Syndiniales), and ellobiopsids, as well as the phagotrophic genus *Oxyrrhis* diverge at the base of the dinoflagellate clade. All later diverging species are known as "core" dinoflagellates and usually are divided into athecate and thecate clades. Thecate dinoflagellates produce cellulosic thecal plates located in the lumens of their alveoli. Most photosynthetic dinoflagellates contain the peridinin plastid, a true cellular organelle that could have been present in the ancestor of "core" dinoflagellates. A number of species later replaced this plastid with other photosynthetic bodies, ranging from "stolen" plastids (termed kleptoplastids) through permanent endosymbionts to fully incorporated plastids (for details, see Table 1). This chapter focuses on three of these photosynthetic entities (1) cryptophyte-derived kleptoplastids in *Dinophysis* (Dinophysiales); (2) permanent, almost complete, diatom endosymbionts in genera like *Durinskia*, *Galeidinium*, and *Kryptoperidinium* (Peridiniales); and (3) haptophyte-derived true plastids found in *Karenia*, *Karlodinium*, and *Takayama* (Gymnodiniales). All these photosynthetic bodies (also known as unusual dinoflagellate plastids) evolved from algae with secondary plastids and, therefore, are characterized as tertiary plastid endosymbioses. At present, it is unclear when the

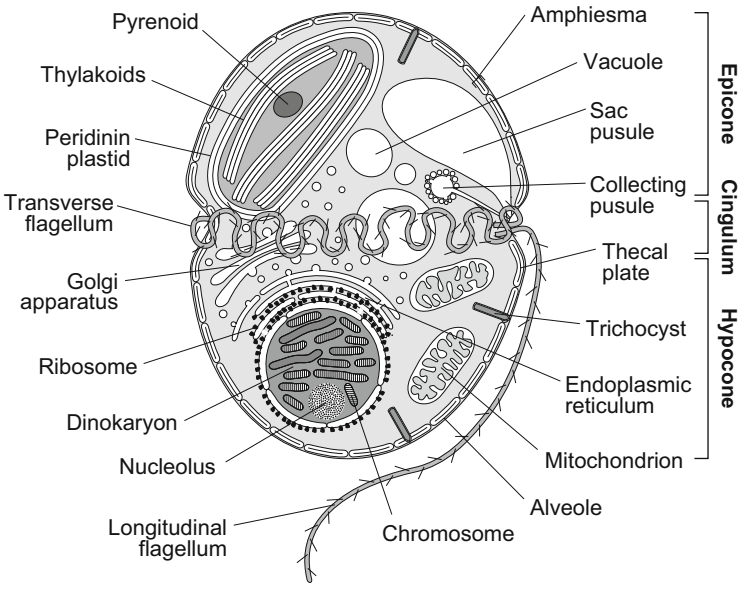


Fig. 2 Ultrastructure of a typical dinoflagellate cell. The cell is divided into the upper epicone and bottom hypocone by the transverse groove or cingulum. The typical dinoflagellate cell has two dissimilar flagella, transverse and longitudinal, which arise from the ventral side of the cell and lie in their respective surface grooves; the transverse flagellum in the cingulum, the longitudinal flagellum in a sulcus that runs perpendicularly to the cingulum. The peculiar transverse flagellum, with its helical axoneme, resembles a wavy ribbon; its beating action makes the cell whirl around its axis as it moves through the water, which gave dinoflagellates their name (Greek *dinos*, "whirling" and Latin *flagellum* "whip, scourge"). The dinoflagellate cell covering, or amphiesma, contains flattened vesicles known as alveoli. In so-called armored dinoflagellates, the alveoli contain cellulosic thecal plates (collectively a theca). Dinoflagellates often have various kinds of extrusomes (e.g., trichocysts), which usually lie perpendicularly to the plasmalemma and play defensive or excretory roles. One of the most unusual features of typical (or "core") dinoflagellates is the unique nucleus, called a dinokaryon, with chromosomes that remain permanently condensed during both interphase and mitosis. A large part of the dinoflagellate cell is filled with a system of vacuoles, including two specialized pusules that mostly are present in heterotrophic, marine species. Pusules are connected to ducts that open near the base of flagella. Dinoflagellate mitochondria have tubular cristae. Most photosynthetic dinoflagellates have the peridinin plastid shown in the figure. It is surrounded by three membranes, with thylakoids stacked in threes, and contains the characteristic carotenoid peridinin

Fig. 1 (continued) peridinin plastid originated. Apicomplexans, perkinsids, and *Oxyrrhis* harbor nonphotosynthetic plastids, whereas a photosynthetic plastid is present in chromerids. Given the distribution of these plastids, it was hypothesized that the peridinin plastid had a common origin with nondinoflagellate plastids, evolving from a red alga via secondary endosymbiosis; however, there also is evidence for an independent origin of the peridinin plastid. For example, it is surrounded by three membranes rather than the four in apicomplexans and perkinsids. Moreover, phylogenetic analyses of plastid genes suggest that the peridinin plastid originated from a haptophyte alga via tertiary endosymbiosis. The peridinin plastid origin may be further clarified by searching for traces of a plastid in ellobiopsids and both MAG groups

using two dissimilar flagella. Particularly noteworthy is the transverse or cingulum-associated flagellum with its ribbon-like structure and helical axoneme (Gaines and Taylor 1985). The most unusual feature of "core" dinoflagellates, however, is the presence of a highly modified nucleus or dinokaryon, which has several distinguishing characteristics; these include (1) permanently condensed chromosomes in a liquid crystalline state, (2) the absence of typical histone nucleosomes and a low protein to DNA ratio (1:10 instead of the 1:1 ratio characteristic of chromatin packed with histones), (3) DNA associated with basic proteins similar to bacterial histone-like proteins (HLPs) that probably were acquired horizontally from bacteria, and (4) extremely large genomes (245 Gb in *Prorocentrum micans* versus 3.2 Gb in the human haploid genome) with even >100 chromosomes (for reviews, see Lin 2011; Wisecaver and Hackett 2011).

In addition to the synapomorphic cellular features noted above, "core" dinoflagellates evolved additional peculiar molecular characters, both organellar and nuclear. First, they have highly reduced plastid genomes organized into numerous 2- to 3-kb plasmid-like minicircles, each containing 0–5 genes (Howe et al. 2008). Second, their mitochondrial genomes exhibit high levels of duplication and recombination and encode only three proteins and two rRNAs (Waller and Jackson 2009). All of these genes occur in multiple copies and frequently are disrupted by noncoding inverted repeats and incomplete gene sequences. Third, frequent and diverse editing of rRNAs and mRNAs occurs in both mitochondria and plastids (Lin et al. 2008). Finally, there is extensive *trans*-splicing of nuclear gene transcripts involving a highly conserved spliced leader sequence, which probably plays a role in regulation of gene expression (Zhang et al. 2007a).

Phylogeny of "Core" Dinoflagellates

The cells of "core" dinoflagellates can be naked (athecate) or armored (thecate) (Fig. 1), the latter having thecal plates composed mainly of cellulose (Kwok and Wong 2003). These plates reside in the cortical alveoli (Pozdnyakov and Skarlato 2012). The different shapes and arrangements of these plates (tabulation) frequently are used as taxonomic characters (see, for example, Fensome et al. 1993). On the basis of such morphological features, "core" dinoflagellates traditionally were divided into various orders, including the thecate Dinophysiales, Gonyaulacales, Peridiniales, Prorocentrales, and Suessiales and the athecate Blastodiniales, Gymnodiniales, and Noctilucales. Molecular phylogenetic studies show that some of these groups are either poly- or paraphyletic (see, for example, Taylor 2004; Moestrup and Daugbjerg 2007; Zhang et al. 2007b); nevertheless, recent analyses by Orr et al. (2012) indicate that athecate forms branch at the base of the dinoflagellate tree, suggesting that intra-alveolar cellulosic plates evolved only once within the "core" dinoflagellates (Fig. 1).

Trophic Strategies in "Core" Dinoflagellates

Mixotrophy: A Trophic Strategy Widespread in Dinoflagellates

"Core" dinoflagellates exhibit a variety of trophic strategies (for reviews, see Stoecker 1999; Schnepf 2004; Jeong et al. 2010; Hansen 2011). Taxa are divided about equally between photoautotrophs and heterotrophs, the latter including saprotrophic, parasitic, and predatory forms. Many dinoflagellate species, however, engage in a mixotrophic lifestyle (Stoecker 1999; Schnepf 2004; Jeong et al. 2010; Hansen 2011). These species represent almost all main dinoflagellate lineages (such as Blastodiniales, Dinophysiales, Gonyaulacales, Gymnodiniales, Noctilucales, Peridiniales, and Prorocentrales) and their eukaryotic prey are diverse, including ciliates, cryptophytes, diatoms, haptophytes, raphidophytes, prasinophytes, and other dinoflagellates (for details, see Stoecker 1999; Schnepf 2004; Jeong et al. 2010; Hansen 2011). Some predominantly heterotrophic species are able to photosynthesize thanks to the presence of temporary photosynthetic bodies (for reviews, see Schnepf 2004; Kim and Archibald 2010; Hansen 2011; Johnson 2011a). The acquisition of phototrophy is profitable for primarily heterotrophic forms under food-limiting conditions, whereas in the case of parasites, it could be important for providing nutrients in periods when hosts are unavailable (e.g., during dispersal stages). Moreover, some mostly photoautotrophic species still can hunt and ingest bacterial and/or eukaryotic prey (Jeong et al. 2010; Hansen 2011). Such prey-capture supplements what is available from inorganic nutrients, which can be limiting, and fosters increased growth rates. Some species that contain permanent plastids cannot grow in the light on standard inorganic nutrient media, and must ingest prey to sustain growth (Jeong et al. 2010; Hansen 2011).

Feeding Mechanisms in Dinoflagellates

Phagocytosis

The ability to engulf whole cells by phagocytosis is widespread in athecate dinoflagellates (e.g., *Gymnodinium*, *Gyrodinium*, and *Noctiluca*) (for reviews, see Hansen and Calado 1999; Schnepf 2004). It also has been observed in some thecate species (e.g., *Ceratium*, *Peridinium*, and *Fragilidium*), but most of these dinoflagellates use more advanced feeding mechanisms, such as myzocytosis and pallium feeding.

Myzocytosis

Dinoflagellates that practice myzocytosis use an appendage (a straw-like structure or feeding tube) to pierce the prey cell and suck out its intracellular content, including whole organelles (for reviews, see Hansen and Calado 1999; Schnepf 2004).

The ingested content is enclosed in one or several food vacuoles within the dinoflagellates, but the prey's plasmalemma is not ingested. It was proposed that the force of suction is generated by lower pressure inside the food vacuole (for a discussion, see Hansen and Calado 1999; Schnepf 2004).

Two types of feeding tubes have been described, the peduncle and the phagopod (for reviews, see Hansen and Calado 1999; Schnepf 2004). The peduncle is a protoplasmic strand of variable length (2–100 μm) and width (1–5 μm), which protrudes from the mid-ventral region of the sulcus and connects the predator to its prey (Hansen and Calado 1999; Schnepf 2004). It is bounded by a single membrane without alveoli and is stiffened by several bands of microtubules. The peduncle usually is retracted in nonfeeding cells; this was found to be the case for example, in *Dinophysis*, *Amphidinium*, *Pfiesteria*, *Gymnodinium*, and *Peridiniopsis* (Hansen and Calado 1999; Schnepf 2004). In histophagous dinoflagellates (*Gymnodinium fungiforme* and *Peridiniopsis berolinensis*), the peduncle also is used to suck fluid contents through existing holes in injured ciliates, other dinoflagellates, and small metazoans (e.g., nematodes and rotifers). Moreover, *Peridiniopsis berolinensis* and *Amphidinium lacustre* are able to take up entire small prey cells through their peduncles (Hansen and Calado 1999; Schnepf 2004).

The phagopod is a hollow cylinder composed of electron-opaque material that does not contain cytoplasmic components such as microtubules or other cytoskeletal elements (for reviews, see Hansen and Calado 1999; Schnepf 2004). In contrast to the peduncle, it forms in an almost antapical position and remains attached to the remnants of the prey after feeding. Dinoflagellates that use this feeding structure include *Amphidinium*, *Gyrodinium*, and *Dissodinium*.

Pallium Feeding

Another unusual mechanism of food acquisition in dinoflagellates is pallium feeding, described so far only in heterotrophic thecate species from the genus *Protoperidinium* and the *Diplopsalis* group (for reviews, see Hansen and Calado 1999; Schnepf 2004). In contrast to phagocytosis and myzocytosis, this kind of feeding proceeds outside of the main cell body (or "extracellularly") in a large pseudopodium called the pallium or feeding veil. When prey is captured with a special peduncle, the pallium emerges from the sulcus area, enlarges and then encloses a single cell or a chain of cells (e.g., diatoms) (Hansen and Calado 1999; Schnepf 2004). After ingestion, the prey resides in a large vacuole where it is digested. Digested material is transported into the cell body, whereas undigested portions (e.g., cell wall) are released into the extracellular environment (Hansen and Calado 1999; Schnepf 2004). After the feeding process is complete, the pallium is retracted.

What Are the Differences Between Endosymbionts and Cell Organelles?

The diversity of trophic strategies in dinoflagellates, and their flexibility in combining and modifying them, led to a remarkable spectrum of photosynthetic bodies (Fig. 3; Table 1). Because these entities reflect different levels of integration with host cells, before describing them, it is important to consider what distinguishes an endosymbiont from a true cell organelle such as a plastid. According to a commonly accepted definition provided by Cavalier-Smith and Lee (1985), an endosymbiont retains all genes for its own proteins and, therefore, its biogenesis does not need to be supported by protein import from the host cell. In contrast, an organelle preserves only a small fraction of its original gene set, and all other required genes reside in the host's nucleus (Cavalier-Smith and Lee 1985). After their expression in the host cytosol, these proteins are imported into a proper cell organelle using targeting signals and import mechanisms in envelope membranes surrounding the organelle.

Based on these considerations, Cavalier-Smith and Lee (1985) proposed two main stages in the evolution of cell organelles. The first is establishment of a permanent endosymbiosis. The prokaryotic and eukaryotic ancestors of cell organelles (including all kinds of plastids) probably were first engulfed as food and digested in food vacuoles (or phagosomes). Consequently, to establish a permanent endosymbiosis, it was necessary to disturb this digestion process, for example, through loss or modification of the phagosomal membrane (Cavalier-Smith and Lee 1985). Only then could a second stage begin that transformed the undigested endosymbiont into a true cellular organelle. This process involved massive gene transfer from the endosymbiont to the host genome over time, along with the evolution of an import apparatus in cell organelle membranes and targeting signals in proteins now encoded in the host's nucleus (Cavalier-Smith and Lee 1985).

Main Kinds of Photosynthetic Bodies in "Core" Dinoflagellates: Peridinin Plastid, Unusual Plastids, and Kleptoplastids

Primary, Secondary, and Tertiary Plastid Endosymbioses in Dinoflagellates

Plastids evolved through multiple levels of endosymbioses (Fig. 3). Primary endosymbionts originate from a photosynthetic cyanobacterium, secondary endosymbionts from a eukaryotic alga with a primary plastid (e.g., a red alga), and tertiary endosymbionts from an alga that already contains a complex, secondary plastid (e.g., a cryptophyte) (for reviews, see Archibald 2009; Keeling 2010). Interestingly, dinoflagellates appear to have experienced all these kinds of

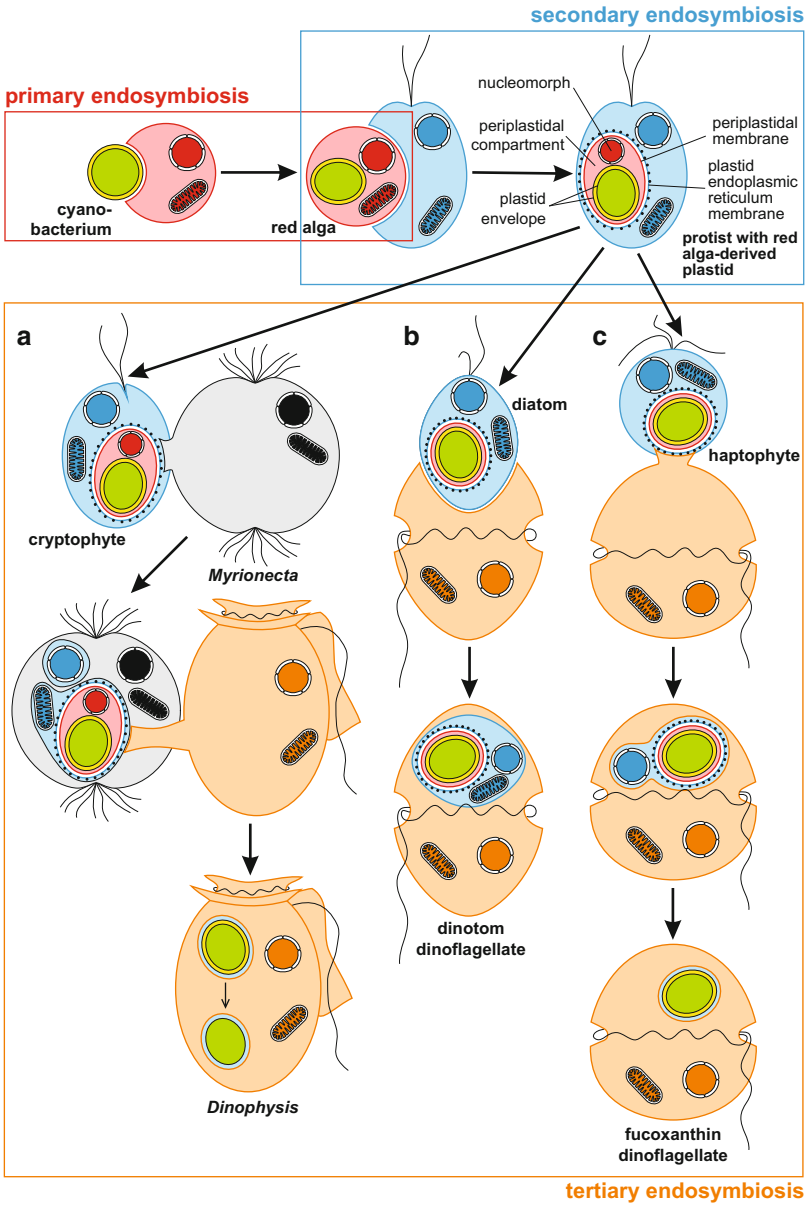


Fig. 3 Three main kinds of tertiary plastid endosymbioses in dinoflagellates. Evolutionary pathways of these unusual plastids can be quite complex. They all started with the primary endosymbiosis, in which a phagotrophic eukaryote engulfed a cyanobacterium that was eventually transformed into a primary plastid with a two-membrane envelope. Primary plastids are characteristic of red algae (as well as green plants and glaucophytes). A red algal cell then was enslaved by a second phagotrophic protist, resulting in a complex, secondary plastid. The ancestral form of this plastid contained a reduced red algal cytoplasm (the periplastidal compartment) and nucleus (or nucleomorph) and was surrounded by four membranes. Its two innermost membranes (the plastid envelope) were derived from the primary plastid, the third (or periplastidal) membrane

endosymbioses and, consequently, contain many types of photosynthetic bodies, including a number of tertiary origin (Table 1). Therefore, the number and variety of photosynthetic entities present, as well as the diverse mechanisms for their acquisition, make the history of endosymbiosis in dinoflagellates exceptional among eukaryotic algae.

Peridinin-Containing Plastid: The Most Typical Dinoflagellate Plastid

Most photosynthetic "core" dinoflagellates contain what is called a peridinin plastid, a true cellular organelle that is surrounded by three membranes (Schnepf and Elbrächter 1999; Nassoury et al. 2003; Patron et al. 2005). It is assumed that its

Fig. 3 (continued) from the red algal plasmalemma, and the outermost from the host's phagosomal membrane. The outermost plastid membrane fused with the host endoplasmic reticulum (ER) resulting in a plastid ER. This red alga-derived plastid commonly is presumed to be present in three main photosynthetic lineages, the cryptophytes, stramenopiles (e.g., diatoms), and haptophytes. The periplastidal compartment and nucleomorph were retained only in cryptophytes. Members of each of these three algal groups have been ingested as tertiary endosymbionts by different dinoflagellate hosts. A: *Dinophysis* species harbor "stolen" plastids (or kleptoplastids) of cryptophyte origin; however, they are acquired indirectly from the ciliate *Myrionecta rubra* via myzocytotic feeding. *Myrionecta rubra* feeds myzocytotically on cryptophyte algae and sequesters their cell organelles within separate compartments containing either the nucleus (kleptokaryon) or plastids and mitochondria. Cryptophyte plastids in *Myrionecta* are complete with four envelope membranes and the nucleomorph. In contrast, *Dinophysis* species ingest only incomplete cryptophyte plastids from *M. rubra*, specifically the compartment contained within the two innermost membranes that corresponds to the original primary plastid. This two-membrane plastid is placed into a food vacuole, meaning *Dinophysis* kleptoplastids initially are surrounded by three membranes. According to one scenario, the middle envelope membrane is digested within the phagosome, resulting in a final two-membrane-bound structure. B: *Kryptoperidinium* and *Durinskia* dinoflagellates engulfed a diatom alga, forming an endosymbiotic consortium known as a dinotom. The endosymbiont's cytosol is separated from the host's by a single membrane. Because almost all of the diatom's intracellular structures are preserved (with the exception of centrioles), it is possible that the surrounding membrane represents the endosymbiont's plasmalemma. Were it derived from the host's phagosomal membrane, the endosymbiont could be at least partially digested. Consequently, it is reasonable to postulate that the diatom was engulfed via phagocytosis and that the phagosomal membrane was lost very quickly. At present, it is unclear whether the endosymbiont imports proteins encoded in the dinoflagellate's nucleus and, therefore, could be considered a true cell organelle. C: The three-membrane-bound fucoxanthin plastid present in the genera *Karenia* and *Karlodinium* probably was obtained via myzocytosis from a haptophyte. After ingestion, the plastid was surrounded by five membranes, but this later was reduced to three, the phagosomal and the two innermost plastid membranes. The endosymbiont's nucleus also must have been ingested during the establishment of this plastid, as a substantial number of haptophyte genes related to plastid functions were transferred to the dinoflagellate nucleus. The fucoxanthin plastid can be characterized as a true cell organelle because it imports host nucleus-encoded proteins

Table 1 Survey of distinct photosynthetic entities in dinoflagellates

Dinoflagellate host	Origin/donor	Selected references
Primary endosymbioses		
Endosymbionts		
Amphisolenia globifera (Dinophysiales)	Cyanobacterium	Lucas (1991)
Dinophysis miles (Dinophysiales)	Cyanobacterium[a]	Qiu et al. (2011)
Sinophysis canaliculata (Dinophysiales)	Cyanobacterium	Escalera et al. (2011)
Secondary endosymbioses		
Plastids		
Lepidodinium chlorophorum, *L. viridae* (Gymnodiniales)	"Core" chlorophyte	Minge et al. (2010)
Endosymbionts		
Noctiluca scintillans (=*N. miliaris*) (Noctulicales)	Prasinophyte *Pedinomonas noctilucae*	Sweeney (1978), Hansen et al. (2004)
Tertiary endosymbioses		
Plastids		
About half of known dinoflagellates	Haptophyte[b]	Schnepf and Elbrächter (1999), Shalchian-Tabrizi et al. (2006)
Karenia brevis, *Karlodinium veneficum*, *Takayama* spp. (Gymnodiniales)	Prymnesiacean haptophyte	Schnepf and Elbrächter (1999), Yoon et al. (2005), Gabrielsen et al. (2011)
Endosymbionts		
Amphisolenia bidentata (Peridiniales)	Pelagophycean stramenopile	Lucas (1991)
Dinothrix paradoxa (Peridiniales)	Diatom	Horiguchi (2006)
Durinskia baltica, *D. capensis*, *Galeidiniium rugatum*, *Kryptoperidinium foliaceum* (Peridiniales)	*Nitzschia*-, *Bacillaria*- or *Cylindrotheca*-like pennate diatom	Inagaki et al. (2000), Imanian and Keeling (2007), Pienaar et al. (2007), Imanian et al. (2010), Figueroa et al. (2009)
Gymnodinium quadrilobatum (Peridiniales)	Diatom	Horiguchi and Pienaar (1994a)
Peridiniopsis penardii, *Peridiniopsis* cf. *kevei* (Peridiniales)	*Discostella*-like centric diatom	Takano et al. (2008), Zhang et al. (2011)
Peridinium quinquecorne (Peridiniales)	*Chaetoceros*-like centric diatom	Horiguchi and Takano (2006)
Podolampas bipes (Dinophysiales)	Stramenopile	Schweikert and Elbrächter (2004)
Kleptoplastids		
Amphidinium latum (Gymnodiniales)	Cryptophyte	Horiguchi and Pienaar (1992)

(continued)

Table 1 (continued)

Dinoflagellate host	Origin/donor	Selected references
Amphidinium poecilochroum (Gymnodiniales)	Cryptophyte *Chroomonas* spp.	Larsen (1988)
Amphidinium wigrense (Gymnodiniales)	Cryptophyte	Wilcox and Wedemayer (1985)
Amylax buxus, *A. triacantha* (Gonyaulacales)	Cryptophyte *Teleaulax amphioxeia*	Koike and Takishita (2008)
Cryptoperidiniopsis spp. (Peridiniales)	Cryptophyte *Storeatula major*	Eriksen et al. (2002)
Dinophysis acuminata, *D. acuta*, *D. caudata*, *D. fortii*, *D. infundibulus*, *D. norvegica*, *D. tripos* (Dinophysiales)	Origin: cryptophyte *Teleaulax amphioxeia*/ *Geminigera criophila*/ *Plagioselmis* sp. Donor: ciliate *Myrionecta rubra*	Minnhagen and Janson (2006), Park et al. (2006, 2008), Nagai et al. (2008), Nishitani et al. (2008a, b, 2010)
Dinophysis miles (Dinophysiales)	Origin: cryptophyte and haptophyte Donor: unknown	Qiu et al. (2011)
Dinophysis mitra (=*Phalacroma mitra*) (Dinophysiales)	Origin: haptophyte, prasinophyte and stramenopile (Bolidophyceae, Dictyochophyceae, Pelagophyceae) Donor: unknown ciliate	Koike et al. (2005), Nishitani et al. (2012)
Gymnodinium aeruginosum (=*G. acidotum*) (Gymnodiniales)	Cryptophyte *Chroomonas* spp.	Wilcox and Wedemayer (1984), Schnepf et al. (1989), Fields and Rhodes (1991)
Gymnodinium gracilentum (Gymnodiniales)	Cryptophyte *Rhodomonas salina*	Skovgaard (1998), Jakobsen et al. (2000)
Gymnodinium myriopyrenoides (Gymnodiniales)	Cryptophyte *Chroomonas* or *Hemiselmis*	Yamaguchi et al. (2011)
"*Kleptodinium*," a dinoflagellate from the Ross Sea (Gymnodiniales)	Haptophyte *Phaeocystis antarctica*	Gast et al. (2007)
Pfiesteria piscida (Peridiniales)	Crypotphyte *Rhodomonas* sp.	Lewitus et al. (1999), Stoecker et al. (2009)

[a]Ecto- or endosymbiont

[b]The peridinin plastid generally is assumed to have a red algal ancestry; although originally hypothesized to be descended from a secondary endosymbiosis in the common ancestor of all "chromalveolates", more recent evidence suggests it was acquired from a haptophyte and, therefore, resulted from a tertiary endosymbiosis (see section "Peridinin-Containing Plastid: The Most Typical Dinoflagellate Plastid")

outermost membrane is derived from the host phagosomal membrane, whereas the two innermost membranes correspond to the original primary plastid envelope (Nassoury et al. 2003; Patron et al. 2005). The photosynthetic apparatus of the peridinin plastid is composed of thylakoids stacked in threes, and contains chlorophylls *a* and *c* along with the characteristic carotenoid peridinin, from which the plastid gets its name (Schnepf and Elbrächter 1999).

The evolutionary origin of the peridinin plastid remains controversial. According to a more traditional view, the plastid evolved from a red alga via secondary endosymbiosis (Cavalier-Smith 1999; Keeling 2009); however, several lines of evidence suggest that it has a tertiary origin involving a haptophyte ancestor (Bodył and Moszczyński 2006; Shalchian-Tabrizi et al. 2006; Bodył et al. 2009a). Whatever evolutionary scenario is correct, nearly all genes were transferred to the host nucleus during the transformation of a eukaryotic alga into the peridinin plastid (Bachvaroff et al. 2004; Hackett et al. 2004b), leaving the plastid genome composed of only highly reduced minicircles (Howe et al. 2008). Each of the transferred genes acquired an N-terminal targeting sequence, usually encoding a signal peptide followed by a transit peptide (class II proteins) (Patron et al. 2005; Patron and Waller 2007; see also Grosche et al. 2013). In one protein group, however, there is an additional hydrophobic domain located after the transit peptide, which probably functions as a stop-transfer domain (class I proteins) (Patron et al. 2005; Patron and Waller 2007). The generally bipartite presequence structure indicates that protein import into the peridinin plastid proceeds in two main stages. First, imported proteins must be targeted to the plastid in transport vesicles via the endomembrane system through the endoplasmic reticulum (ER) and/or Golgi apparatus (Nassoury et al. 2003; Patron et al. 2005; Grosche et al. 2013). After fusing with the outermost plastid membrane, these vesicles release their proteins into the intermembrane space, from where they are translocated across the two innermost membranes. The initial stage depends on the signal peptide and the Sec translocon, whereas movement across the inner two membranes involves the transit peptide and Toc- and Tic-like translocons (Nassoury et al. 2003; Patron et al. 2005; Grosche et al. 2013).

At present, it is unclear whether the dinoflagellate ancestor contained a peridinin plastid although several hypotheses have been put forward (see, for example, Cavalier-Smith 1999; Bodył and Moszczyński 2006; Janouskovec et al. 2010). If the dinoflagellate ancestor did indeed harbor such a plastid, evidence of this organelle should be found in early branching lineages such as ellobiopsids, MAG I, MAG II/Syndiniales, and *Oxyrrhis* (Fig. 1). Although no plastid has been identified in *Oxyrrhis*, it does contain genes encoding proteins with bipartite presequences, which are characteristic of eukaryotic alga-derived plastids (Slamovits and Keeling 2008). As yet, however, there is no evidence for current or historical plastid residence in the remaining early-branching lineages. Thus, it is possible that the peridinin plastid was acquired within dinoflagellates, by a later diverging lineage (Fig. 1).

Unusual Dinoflagellate Plastids and Their Evolutionary Pathways

In addition to the peridinin plastid, dinoflagellates contain other types of photosynthetic bodies (Table 1). Generally, four main kinds of such unusual dinoflagellate plastids are distinguished: (1) a green algal plastid in *Lepidodinium chlorophorum*

and *L. viride* (e.g., Minge et al. 2010), (2) a cryptophyte-derived plastid in the genus *Dinophysis* (e.g., Hackett et al. 2003), (3) a stramenopile plastid acquired by *Kryptoperidinium foliaceum* and *Durinskia baltica* (e.g., Inagaki et al. 2000), and (4) a haptophyte-derived plastid in representatives of the genera *Karenia* and *Karlodinium* (e.g., Tengs et al. 2000).

The *Lepidodinium* green plastid has a secondary origin and appears to represent a true cellular organelle, as indicated by its permanent residence in the host cell and the import of numerous host nucleus-encoded proteins (Minge et al. 2010). The three remaining types of unusual dinoflagellate plastids evolved via tertiary endosymbioses (Fig. 3) and are the main subject of this chapter. Although these photosynthetic bodies usually are called plastids, available data demonstrate that at least one of them, the entity in *Dinophysis*, cannot be considered a true organelle (see, for example, Wisecaver and Hackett 2010). Beginning in section "Cryptophyte-Derived Photosynthetic Bodies of *Dinophysis* species: True Plastids or Kleptoplastids?," we discuss specific details of each of the unusual tertiary photosynthetic entities in "core" dinoflagellates.

In 1993, Schnepf first proposed the interesting idea that the acquisition of unusual dinoflagellate photosynthetic bodies was facilitated by the prior existence of the peridinin plastid (Schnepf 1993). This replacement model was further developed by Häuber et al. (1994), Bodył (1999), and Cavalier-Smith (1999). The main idea of a replacement model is that the newly acquired photosynthetic endosymbiont could make use of preexisting host nucleus-residing genes and protein-targeting machinery from the original peridinin plastid. Thus, a new algal endosymbiont could be transformed into a true organelle much more easily than in the case of a previously aplastidal host. The process would require neither massive gene transfer nor evolution of new targeting signals and import system.

Kleptoplastids: A Peculiar Kind of Photosynthetic Bodies in Dinoflagellates

"Core" dinoflagellates not only harbor permanent photosynthetic endosymbionts and true plastids but also transiently sequestered plastids (Table 1). These "stolen" organelles are known as kleptoplastids (for reviews, see Schnepf 2004; Kim and Archibald 2010; Johnson 2011a). "Core" dinoflagellates usually engulf kleptoplastids via myzocytosis along with other prey cellular components (e.g., nucleus), but nonplastid cell constituents are degraded quickly. Consequently, kleptoplastids are active for up to 1 month in most cases, and sometimes only for several days (Schnepf 2004; Kim and Archibald 2010; Johnson 2011a). The "stolen" plastids provide their heterotrophic hosts with photosynthetic metabolites, and host cells evolve mechanisms that extend the time kleptoplastids can survive in the endomembrane system or cytosol. Kleptoplastids also are found in nondinoflagellate protists, including ciliates, foraminiferans, centrohelid heliozoans, and in animals (e.g., Sacoglossan sea slugs)

(for details, see Stoecker et al. 2009; Johnson 2011a; Rumpho et al. 2011; see also Wägele and Martin 2013).

As noted above, most kleptoplastids are short-lived photosynthetic entities. An example of such kleptoplastids are those found in *Gymnodinium aeruginosum*, also known as *G. acidotum* (Wilcox and Wedemayer 1984; Schnepf et al. 1989; Farmer and Roberts 1990). Only a single cryptophyte-derived kleptoplastid is present per *G. aeruginosum* cell. In addition to the plastid, this dinoflagellate ingests the cryptophyte's mitochondria and endomembrane system, although its nucleus most often is absent (in 70–90 % of observed cases) (Wilcox and Wedemayer 1984; Schnepf et al. 1989; Farmer and Roberts 1990). The cryptophyte kleptoplastid and other cell structures are kept in a vacuole that spreads throughout the host cell. This enlargement appears to indicate an advanced level of integration between these kleptoplastids and their host cells. Although it was initially suggested that *G. aeruginosum* contains an endosymbiont, rather than a kleptoplastid (Wilcox and Wedemayer 1984; Farmer and Roberts 1990), Fields and Rhodes (1991) showed that the dinoflagellate fundamentally is a colorless species that feeds on *Chroomonas* cryptophytes. The ingested cryptophyte plastids are maintained for about 12 days.

There are dinoflagellates, however, that can maintain active kleptoplastids for months. An abundant but not formally described species from the Ross Sea (Antarctica) feeds on the bloom-forming haptophyte *Phaeocystis antarctica* (Gast et al. 2007). Both the Ross Sea dinoflagellate and its haptophyte prey are found together in seawater and slush (a mixture of seawater and ice) habitats. The dinoflagellate acquires up to 20 kleptoplastids from *P. antarctica*, but the number decreases in the absence of additional plastid capture from the haptophyte (Gast et al. 2007). Nevertheless, the remaining plastids can be maintained for as long as 5–8 months, which is highly unusual for kleptoplastids. The Ross Sea dinoflagellate appears to be fully dependent on a continuous kleptoplastid endosymbiosis with *P. antarctica*; although it remains viable, it cannot grow in the light without being fed *P. antarctica*, nor in darkness even when the haptophyte prey is plentiful (Gast et al. 2007). In addition to haptophyte-derived plastids, a second, nondinokaryotic nucleus was observed in the Ross Sea dinoflagellate, although its origin has not been determined. Nevertheless, Gast et al. (2007) hypothesized that the very long maintenance of kleptoplastids could be supported by the second nucleus possibly also derived from *P. antarctica* cells.

Cryptophyte-Derived Photosynthetic Bodies of *Dinophysis* Species: True Plastids or Kleptoplastids?

General Characteristic of Dinophysis

The dinoflagellate genus *Dinophysis* includes over 100 species (morphotypes) and is broadly distributed in coastal and oceanic waters (for a review, see Reguera

et al. 2012). *Dinophysis* cells show strong lateral compression and are enclosed in a theca with a characteristic funnel-shaped fan; the theca consists of a few large plates and a series of small platelets (Fensome et al. 1993; Reguera et al. 2012). Thecal morphology exhibits great variation so that species of *Dinophysis* in a given biogeographic region can show a continuum of shapes and sizes resulting from their polymorphic lifecycles. This causes uncertainty in the identification of environmentally collected samples to particular species.

The genus *Dinophysis* is classified in the order Dinophysales together with about 15 other genera, including *Amphisolenia*, *Ornithocercus*, *Histioneis*, *Citharistes*, and *Phalacroma* (Jensen and Daugbjerg 2009; Gómez et al. 2011; Gómez 2012b). Most dinophysalian dinoflagellates are predominantly heterotrophic, feeding on other protists; however, reversions to photosynthesis happened in several genera through endosymbiotic acquisitions of distinct eukaryotic algae (Fig. 3A; Table 1). For example, *Amphisolenia bidentata*, which branches at the base of the dinophysoid tree, harbors a stramenopile alga from the class Pelagophyceae (Schweikert and Elbrächter 2004). A number of apochlorotic species were identified in *Dinophysis* (*D. hastata*, *D. odiosa*, *D. monacantha*, *D. pusilla*, and *Dinophysis cf. acutissima*) and others that contain cryptophyte-derived chromatophores (*D. acuminata*, *D. acuta*, *D. fortii*, *D. infundibulus*, *D. caudata*, *D. tripos*, *D. norvegica*, and *D. miles*) (see Qiu et al. 2011; Gómez et al. 2011 and references therein). These two groups of *Dinophysis* species occur in separate clades in phylogenetic analyses, indicating that there is some genetically based predisposition to plastid acquisition in the genus (Gómez et al. 2011). A haptophyte kleptoplastid also was found in *Phalacroma mitra*, a sister genus to *Dinophysis* (Koike et al. 2005).

Dinophysoid dinoflagellates also are known to have cyanobacterial symbionts (Table 1). *Ornithocercus*, *Histioneis*, *Parahistioneis*, and *Citharistes* all developed ectosymbiotic relationships with cyanobacteria that reside in special extracellular cavities and cones (Taylor 1976; Hallegraeff and Lucas 1988; Foster et al. 2006a, b; Tarangkoon et al. 2007). Intracellular cyanobacterial symbionts were found in the genus *Amphisolenia*, with its characteristically elongated host cells (Lucas 1991), and in the benthonic species *Sinophysis canaliculata* (Escalera et al. 2011). Recent studies identified a nitrogenase (*nifH*) gene of proteobacterial origin in *Ornithocercus* and *Amphisolenia*; together with microscopic observations, this suggests the presence of additional N_2-fixing heterotrophic symbionts (Farnelid et al. 2010). Cyanobacterial and other bacterial symbionts found probably provide fixed carbon and nitrogen to their hosts (Foster et al. 2006a, b; Tarangkoon et al. 2007; Farnelid et al. 2010).

Some *Dinophysis* species pose major economic and health concerns during seasonal blooms (red tides) that can produce toxins (for a review, see Reguera et al. 2012). They can synthesize two kinds of lipophilic toxins: okadaic acid and its dinophysistoxin derivatives, and pectenotoxins. Interestingly, particular strains are able to produce only one group of toxins. These harmful algal blooms cause major economic losses in the shellfish industry, especially in Europe and Japan. Moreover, the toxins accumulate in shellfish and can cause unpleasant gastrointestinal symptoms known as diarrhetic shellfish poisoning (DSP) in human consumers (Schnepf and Elbrächter 1988; Marasigan et al. 2001; Nielsen et al. 2012).

Evidence for Cryptophyte Origin of Dinophysis *Chromatophores*

The envelope of *Dinophysis* photosynthetic bodies (or chromatophores) is composed of only two membranes (Fig. 3A). Thylakoids are found in pairs or stacks of three with a characteristic electron-dense lumen (Schnepf and Elbrächter 1988; Lucas and Vesk 1990; Garcia-Cuetos et al. 2010). Each chromatophore terminates with a pyrenoid. Because the pyrenoids group together into the so-called compound pyrenoid, *Dinophysis* chromatophores form a characteristic stellate structure (Schnepf and Elbrächter 1988; Garcia-Cuetos et al. 2010).

Given the presence of only two envelope membranes, Cavalier-Smith (1993b) originally suggested that *Dinophysis* chromatophores are derived from the same cyanobacterial endosymbiont as the plastids of glaucophytes, red algae, and green plants (see Löffelhardt 2013) and, therefore, represent one additional lineage of ancient, primary plastids. He even created a separate class, the Bilidinea, for *Dinophysis* species. Thylakoid architecture strongly resembles that of cryptophytes, however, which pointed to a tertiary cryptophyte origin of *Dinophysis* chromatophores (Schnepf and Elbrächter 1988; Lucas and Vesk 1990). In a further support of this view, Vesk et al. (1996) demonstrated that phycoerythrin is located in the thylakoid lumen of *D. acuminata* and *D. fortii* chromatophores, which otherwise is observed only in cryptophyte plastids. Moreover, the most abundant carotenoid pigment in *Dinophysis* is alloxanthin, also characteristic of cryptophytes (Meyer-Harms and Pollehne 1998; see also Hewes et al. 1998). The controversy over the evolutionary origin of *Dinophysis* chromatophores was settled by phylogenetic analyses of their genes, which demonstrated unequivocally that they are derived from cryptophyte algae (see, for example, Takishita et al. 2002; Janson and Granéli 2003; Janson 2004; Nishitani et al. 2010)

Indirect Myrionecta rubra*-Mediated Route of Cryptophyte Plastids into* Dinophysis *Cells*

The data discussed above clearly demonstrate that photosynthetic *Dinophysis* species harbor cryptophyte-derived chromatophores (Fig. 3A). Because these dinoflagellates often contain food vacuoles (Jacobson and Andersen 1994; Koike et al. 2000), it was expected that their growth in cultures could be strictly dependent on the presence of prey organisms. Unexpectedly, all attempts to cultivate photosynthetic *Dinophysis* species with various cryptophytes and other potential prey (e.g., diatoms) were unsuccessful for many years (see, for example, Sampayo 1993; Nishitani et al. 2003). Considering these difficulties, Janson (2004) proposed that *Dinophysis* species feed not on cryptophytes directly, but on the ciliate *Myrionecta rubra* (=*Mesodinium rubrum*) (Fig. 3A). *Myrionecta rubra* is a mixotrophic red tide-causing species that feeds upon cryptophyte algae and maintains them as

kleptoplastids (for a review, see Johnson 2011b). The Janson's hypothesis was completely confirmed 2 years later by Park et al. (2006), with the first successful establishment of a long-term culture of *D. acuminata*. Park et al. (2006) also observed that *D. acuminata* ingests the cryptophyte plastids residing in *M. rubra* myzocytotically via a peduncle. In subsequent years, long-term cultures were established in the same manner for other *Dinophysis* species as well, including *D. fortii* (Nagai et al. 2008), *D. caudata* (Nishitani et al. 2008a), and *D. infundibulus* (Nishitani et al. 2008b). These experiments clearly showed that *Dinophysis* species receive their chromatophores indirectly by feeding on *M. rubra* that contains intracellular cryptophytes.

The evolutionary story of *Dinophysis* chromatophores is even more complex owing to the peculiar ability of *M. rubra* cells to sequester cryptophyte cell organelles and place them into distinct cellular compartments (Fig. 3A). One of these compartments contains cryptophyte plastids and mitochondria that form characteristic plastid–mitochondrion complexes (Oakley and Taylor 1978; Gustafson et al. 2000; Johnson et al. 2006). In contrast to *Dinophysis* chromatophores, the cryptophyte plastids engulfed by *M. rubra* still have their original structure, complete with four envelope membranes and the nucleomorph. The cryptophyte nucleus resides in a separate compartment from the plastid–mitochondrion complexes within *M. rubra* (Oakley and Taylor 1978; Johnson et al. 2007). Because nuclei and plastids are integrally connected within intact cryptophyte cells, *M. rubra* must separate them during the process of ingesting cryptophytes. The sequestered cryptophyte nuclei, termed kleptokaryons, remain transcriptionally active for several weeks and are regularly supplemented by fresh nuclei ingested from new cryptophyte prey (Johnson et al. 2007). Interestingly, high expression levels for genes encoding plastid proteins were reported not only in the kleptokaryon but also in the nucleomorph and the cryptophyte plastid. The whole process of sequestration and maintenance of cryptophyte nuclei is referred to as karyoklepty (Johnson et al. 2007).

Nature of the Cryptophyte-Derived Chromatophores in Dinophysis *Species: True Plastids or Kleptoplastids?*

Why *Dinophysis* Species Feed on *Myrionecta rubra* Cells?

The strict dependence of *Dinophysis* cells on *M. rubra* prey, as demonstrated for the first time by Park et al. (2006), suggested that their chromatophores are kleptoplastids, meaning they would need to be supplemented regularly with newly ingested cryptophyte plastids. At that time, however, and for several years after, the alternative evolutionary scenario that these dinoflagellates harbor some permanent plastids (see, for example, Park et al. 2010) could not be excluded. In such a case, they would feed on *M. rubra* for reasons other than to acquire

additional temporary kleptoplastids. For example, *Dinophysis* cells could require growth factors to ensure efficient growth of host cells and/or their permanent plastids (Garcia-Cuetos et al. 2010; see also Skovgaard 2000). Consequently, the acquisition of new kleptoplastids would be only a by-product of the feeding process.

Do Ultrastructural Data Support a True Plastidic Nature of *Dinophysis* Chromatophores?

Ultrastructural data usually are adduced to support a true plastidic nature of *Dinophysis* chromatophores. The basic argument points to the presence of only two envelope membranes in these chromatophores, that is, they lack the additional envelope membranes present in cryptophyte plastids, as well as the cryptophyte's nucleomorph, nucleus, and cytoplasm (Schnepf and Elbrächter 1988; Lucas and Vesk 1990). Because no vestiges of these structures ever were found, it was argued that *Dinophysis* chromatophores must be true plastids (see, for example, Schnepf and Elbrächter 1999; Schnepf 2004); however, these observations also are consistent with a kleptoplastidic scenario. All cryptophyte constituents, with the exception of the inner plastid compartment surrounded by two envelope membranes, could quickly be digested after ingestion; alternatively, only the inner plastid compartment is taken up in the first place. This latter model is strongly supported by Kim et al. (2012), who found that *Dinophysis* cells ingest only the inner plastid compartment from the cryptophyte plastids residing in *M. rubra* cells. The incomplete cryptophyte plastids first are placed into a central food vacuole and then they migrate to the cell periphery.

A new argument favoring a true plastidic scenario based on ultrastructure was advanced by Garcia-Cuetos et al. (2010), who pointed out two important ultrastructural differences between the cryptophyte plastids in *T. amphioxeia* and *M. rubra*, and those in *D. acuminata*. First, pyrenoids are lateral and single in *Teleaulax* and *Myrionecta*, but terminal and compound in *Dinophysis*. Second, thylakoids are arranged in triplets in the cryptophyte and the ciliate, whereas they occur in pairs in the dinoflagellate. Garcia-Cuetos et al. (2010) argued that such important ultrastructural differences, along with different number of envelope membranes and the absence or presence of a nucleomorph, could result only from the complete transformation of the cryptophyte plastids into true cell organelles in the dinoflagellate. Recent ultrastructural studies by Kim et al. (2012), however, demonstrate clearly that all these changes occur within *Dinophysis* cells, after migration of ingested incomplete cryptophyte plastids from the central phagosome to the cell periphery. It also was observed that chromatophores group together to form their characteristic stellate structure in *Dinophysis* after they reach their final destination.

Paradoxically, although Garcia-Cuetos et al. (2010) used ultrastructural analyses to support a true plastidic scenario for *Dinophysis* chromatophores, their studies provided evidence for the alternative kleptoplastidic model. One of their most important findings was the identification of fragments of a third membrane between

the outer and inner envelope membranes in *D. acuminata* chromatophores. Because they reside in close proximity to the inner membrane, Garcia-Cuetos et al. (2010) hypothesized that these vestigial membranes correspond to the outer membrane of primary plastids, that is, the original cyanobacterial outer membrane. These authors concluded that these ultrastructural data represent evidence for a true plastidic nature of *Dinophysis* chromatophores. The alternative kleptoplastid explanation is much more probable, however, because the outer membrane of *Dinophysis* chromatophores likely is derived from the host's phagosomal membrane, which remains intact from the ingestion process. The incomplete cryptophyte plastids (devoid of their two outermost membranes, that is, the plastid ER membrane and the periplastidal membrane) are placed initially into the central food vacuole, where the outer plastid membrane probably is partially digested (Kim et al. 2012).

Molecular Data Provide Strong Support for Kleptoplastidy in *Dinophysis* Species

Analyses of 16S rDNA and *psbA* plastid genes demonstrated that their sequences are identical among distinct photosynthetic *Dinophysis* species (Takishita et al. 2002; Janson and Granéli 2003). This also is true between cells of the same species living in widely separated geographic areas, such as the Japanese Sea and the Baltic Sea. It was also confirmed for the plastid 16S rDNA sequences that they resided within chromatophores, not food vacuoles (Takishita et al. 2002). All these data clearly support the kleptoplastidic nature of *Dinophysis* chromatophores. If the chromatophores were true plastids, there should be at least some divergence in their DNA sequences, as was shown for organelle genomes in both fucoxanthin and peridinin dinoflagellate plastids (Zhang et al. 2000; Shalchian-Tabrizi et al. 2006). The lack of differences in plastid sequences among *Dinophysis* species clearly contrasts with divergences reported in their nuclear 18S rDNA (Takishita et al. 2002). On the other hand, if *Dinophysis* chromatophores are actually kleptoplastids, then their exact source should be found among currently living organisms. Initial phylogenetic analyses were inconclusive (Takishita et al. 2002; Janson and Granéli 2003), but later studies demonstrated that most *Dinophysis* plastid sequences are identical to those in the cryptophyte *Teleaulax amphioxeia* (Janson 2004; Minnhagen and Janson 2006; Park et al. 2008; Garcia-Cuetos et al. 2010; Nishitani et al. 2010), indicating clearly that this cryptophyte species is the source of their kleptoplastids.

The kleptoplastidic nature of *Dinophysis* chromatophores has found support in additional studies. First, analysis of additional plastid molecular markers in *Dinophysis* spp., including the more variable Intergenic Transcribed Spacer (ITS) region, confirmed the lack of sequence variation among plastids in distinct *Dinophysis* strains or species, and between *Dinophysis* and *Teleaulax* (Minnhagen and Janson 2006). Second, by comparing amounts of plastid DNA in dividing and nondividing *D. norvegica* cells, Minnhagen et al. (2008) found that there was no replication of plastid genomes, indicating that *Dinophysis* chromatophores do not

undergo division and, therefore, cannot be permanent structures. Third, sequence identity of plastid genes not only was demonstrated between *T. amphioxeia* and *Dinophysis* species but also in the ciliate *M. rubra* (see, for example, Park et al. 2008; Garcia-Cuetos et al. 2010; Nishitani et al. 2010), providing strong support for an indirect *M. rubra*-mediated route of cryptophyte plastids into *Dinophysis* cells.

Genomic Evidence for Kleptoplastidy in *Dinophysis* Species

Any remaining debate over the kleptoplastidic versus true plastidic nature of *Dinophysis* chromatophores appears to have been resolved by analyses of ESTs (expressed sequence tags) from *D. acuminata* (Wisecaver and Hackett 2010). Plastids generally require about 2,000 proteins to fulfill their functions (Richly and Leister 2004; van Wijk 2004), but their genomes encode only 100–200 proteins (Green 2011). Because all remaining plastid proteins are encoded in the host nuclear genome, plastid biogenesis is strictly dependent on the import of hundreds of nuclear-encoded proteins (for reviews, see Inaba and Schnell 2008; Agrawal and Striepen 2010; see also Grosche et al. 2013 and Sommer and Schleiff 2013). The recently published complete genomic sequence of the cryptophyte *Guillardia theta* indicates that its complex plastids import about 3,400 nuclear-encoded proteins, most of them targeted to the periplastidal compartment and the nucleomorph, and some 800 directly to the inner plastid compartment itself (Curtis et al. 2012). Consequently, if cryptophyte-derived chromatophores in *Dinophysis* were true plastids, the host nuclei would have to encode hundreds of plastid-targeted proteins that formerly were encoded in the cryptophyte nucleus and nucleomorph. Wisecaver and Hackett (2010) found only five plastid sequences in the *D. acuminata* EST database, however, (1) the photosystem II-associated PsbM, (2) PsbU, another subunit of photosystem II, (3) light harvesting protein LI818, (4) ferredoxin, and (5) triose-phosphate transporter (TPT) (see also Kim and Archibald 2010). Thus, given the absence of hundreds of essential plastid genes from the *D. acuminata* nuclear genome, or anywhere else within the dinoflagellate cell, these photosynthetic bodies must be kleptoplastids.

Protein Import into Dinophysis *Kleptoplastids and the Origin of Their Envelope Membranes*

Distinct Post- and Cotranslational Import Routes of Nuclear-Encoded Plastid Proteins

Each of the five plastid proteins encoded in the *D. acuminata* nuclear genome is equipped with an N-terminal targeting signal (Fig. 4A); PsbM, PsbU, LI818, and TPT carry one domain signals resembling classical plastid transit peptides, whereas

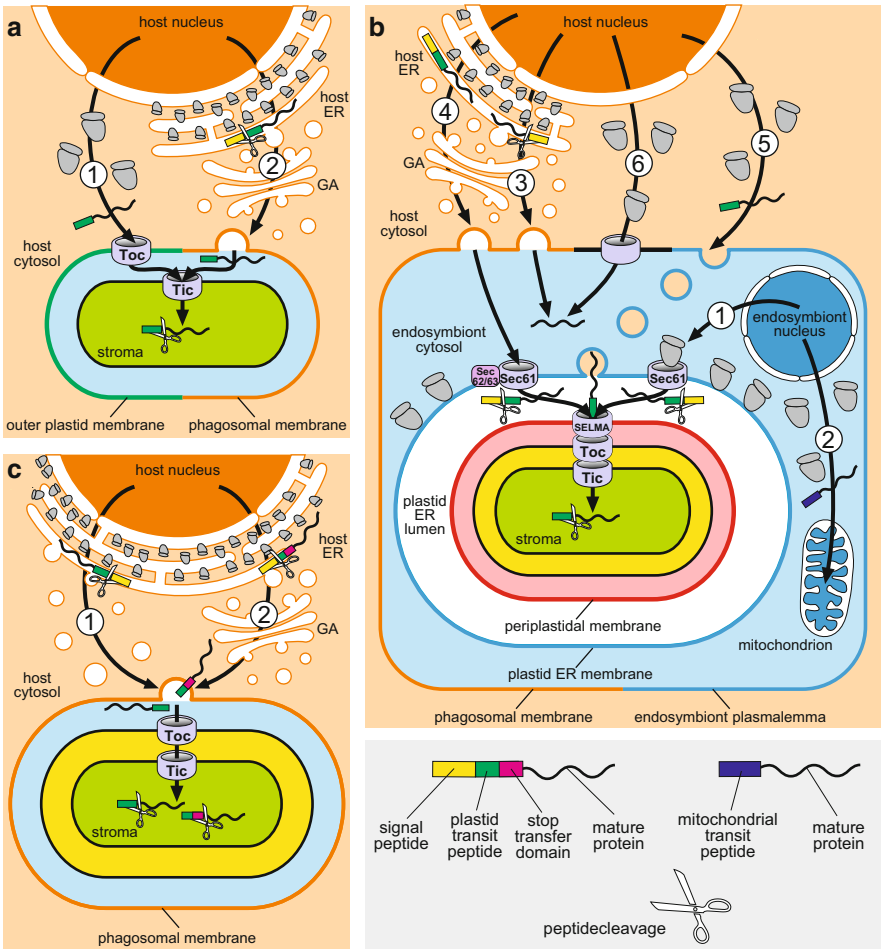


Fig. 4 Mechanisms of protein import into unusual dinoflagellate plastids of tertiary origin. A: Two membranes surround cryptophyte-derived kleptoplastids of *Dinophysis*. Combined with the presence of nuclear-encoded proteins with N-terminal transit peptides, this suggests that protein import into the stroma proceeds posttranslationally via canonical Toc and Tic translocons (1). Consequently, the two membranes likely correspond to the original primary plastid envelope; however, the discovery of a protein with a bipartite presequence, composed of a signal peptide followed by a transit peptide, suggests that cotranslational import via the endomembrane system also occurs, involving the endoplasmic reticulum (ER) and/or Golgi apparatus (GA) (2). This kind of transport is more characteristic of a phagosomal origin of the outer membrane; therefore, the outer membrane of *Dinophysis* kleptoplastids probably is chimeric in nature. B: The diatom endosymbiont of dinotoms is surrounded by a single membrane. The unreduced diatom nucleus still is present and probably encodes most of the proteins required by the plastid. It is likely that plastid-protein import is similar to that of free-living diatoms. Consequently, plastid-targeted proteins equipped with a bipartite presequence, comprising a signal peptide and a transit peptide, would be cotranslationally translocated across the plastid ER membrane using the Sec61 translocon. After cleavage of their signal peptide, the proteins would be imported through the periplastidal membrane using ERAD-like machinery (SELMA), whereas the two innermost plastid membranes would be crossed via Toc and Tic translocons (1). Endosymbiont mitochondria most likely are supplied by endosymbiont nucleus-encoded proteins containing mitochondrial

ferredoxin is equipped with a bipartite targeting signal composed of a signal peptide followed by a transit peptide (Wisecaver and Hackett 2010). The presence of nondegenerated targeting signals in all these proteins strongly suggests that they actually are imported into *D. acuminata* kleptoplastids.

The fact that *D. acuminata* plastid-targeted proteins carry both one- and two-domain targeting signals suggests that there are at least two protein import routes into their kleptoplastids (Fig. 4A). Based on their transit peptides, PsbM, PsbU, LI818, and TPT would be imported posttranslationally through Toc and Tic translocons embedded in their outer and inner envelope membranes, respectively (Inaba and Schnell 2008; Sommer and Schleiff 2013). In contrast, with its bipartite leader peptide, ferredoxin would be targeted via the host endomembrane system, the signal peptide enabling cotranslational passage into the ER, from where transport vesicles could deliver it to the kleptoplastid surface (Agrawal and Striepen

Fig. 4 (continued) transit peptides (2). Given the disorganized division of the endosymbiont nucleus, however, during which essential genes could be lost, it is possible that the diatom endosymbiont imports at least some proteins encoded in the host nucleus. These proteins could be targeted to distinct compartments within the diatom endosymbiont, such as its cytosol, plastids, mitochondria, and nucleus. Assuming the membrane surrounding the endosymbiont is phagosomal in origin, endosymbiont-targeted proteins destined to its cytosol (3) and plastid (4) would have to carry signal peptides. These peptides would result in translocation into the host endomembrane system, from where transport vesicles would deliver them to the endosymbiont's surface, fuse with the surrounding endosymbiont membrane, and release the proteins into the endosymbiont's cytosol. Cytosolic proteins would remain there, whereas plastid proteins would be targeted to the plastid. Unlike in cytosol-targeted proteins, signal peptides of plastid proteins would not be removed in the host ER lumen because of the need for further translocation across the endosymbiont's plastid ER membrane. They would be recognized by receptor proteins (homologous to the yeast and mammalian Sec62/63 complex) associated with the Sec61 translocation channel in the plastid ER membrane. The signal peptides finally would be cut off in the plastid ER lumen and proteins, now with exposed transit peptides, would be translocated into the stroma using SELMA as well as Toc and Tic translocons. Alternatively, if the endosymbiont is surrounded by a plasmalemma that still retains pinocytotic properties, plastid-destined proteins could be taken up by pinocytosis (5). In this case, proteins with a transit peptide but no signal peptide would be released into the plastid ER lumen after pinocytotic vesicles fused with the outermost plastid membrane and would be imported further via the same SELMA–Toc–Tic pathway as described above. Regardless of the provenance of the surrounding endosymbiotic membrane, diatom-targeted proteins could be translocated across it through a protein-conducting channel (6). C: The majority of proteins imported into fucoxanthin plastids carry bipartite N-terminal leader sequences consisting of a signal peptide followed by a transit peptide (class II proteins). The signal peptide results in cotranslational import into the endomembrane system, from where proteins possibly are targeted to the outermost plastid membrane (of phagosomal origin) in vesicles (1). An additional hydrophobic region downstream of the transit peptide was identified in some proteins (class I proteins), which probably functions as a stop transfer domain, anchoring plastid proteins to the ER membrane as they are delivered to the plastid (2). Based on plastid protein import in peridinin dinoflagellates and euglenids, it could be hypothesized that class II proteins are targeted only via the ER, whereas class I proteins move through both the ER and the GA. After a transport vesicle fuses with the outermost plastid membrane, members of both protein classes would be further translocated into the stroma by Toc and Tic translocons that recognize their transit peptides

2010; Grosche et al. 2013). After the vesicle fuses with the outer envelope membrane, the protein would be released into the intermembrane space and could cross the inner envelope membrane through the Tic translocon using the second, transit peptide-like domain (Fig. 4A).

The use of multiple import systems into *D. acuminata* kleptoplastids would not be unprecedented. In higher plant plastids, most proteins are imported posttranslationally using the Toc–Tic-based and transit peptide-dependent pathway; however, some proteins carry signal peptides and are targeted via the ER and/or Golgi apparatus (for reviews, see Inaba and Schnell 2008; Bodył et al. 2009b).

Is the Outer Membrane of *Dinophysis* Kleptoplastids Derived from the Cryptophyte Plastid or the Host Phagosome?

The presence of transit peptides in almost all plastid proteins identified in *D. acuminata* by Wisecaver and Hackett (2010) suggests that the two membranes surrounding *Dinophysis* kleptoplastids correspond to the two membrane envelope of primary plastids (Fig. 4A). Consequently, these membranes would contain Toc and Tic translocons, enabling posttranslational import of host nucleus-encoded proteins (Inaba and Schnell 2008; Sommer and Schleiff 2013). In accordance with the hypothesis, recent studies by Kim et al. (2012) suggest that the two-membrane-bound cryptophyte plastids, which reside in the central food vacuole when ingested, might be released directly into the *Dinophysis* cytosol, perhaps with the help of surrounding vesicles within the phagosome. These interesting and surprising new data need confirmation via further research; thus, alternative scenarios must be considered.

As discussed in the section "Why *Dinophysis* species Feed on *Myrionecta rubra* Cells?," it appears most likely that the outer envelope membrane of *Dinophysis* kleptoplastids is derived from the host phagosomal membrane, with the inner membrane corresponding to the innermost membrane in all types of plastids (Fig. 4A). A phagosomal genesis of the outer membrane is supported by the fact that cryptophyte plastids ingested from *M. rubra* first are placed into the central food vacuole (Kim et al. 2012). Moreover, the phagosomal hypothesis explains the remnants of a third membrane located between the outer and inner envelope membranes of *D. acuminata* kleptoplastids reported by Garcia-Cuetos et al. (2010). It is possible that the central food vacuole is divided over time into smaller vacuoles, each containing a single cryptophyte-derived kleptoplastid. To prevent digestion of the plastid, their surrounding membrane could be modified impeding fusion with prelysosomal vesicles. Through further modifications the membrane could acquire different host- and/or endosymbiont-derived transporters, enabling efficient exchange of distinct compounds between the kleptoplastids and the host.

If the outer membrane of *Dinophysis* kleptoplastids is actually derived from the host phagosomal membrane, and originally devoid of any plastid translocons, how could it have acquired a machinery that interacts with transit peptides to move proteins across the membrane? It is possible that channel (and also receptor)

components of the translocons resembling those of primary plastids, such as mitochondrial Tom and Tim translocons, are inserted into the modified phagosomal membrane surrounding *Dinophysis* kleptoplastids, enabling import of transit peptide-carrying plastid proteins. Such a scenario would be analogous to what occurs in some higher plant plastid proteins that are imported through the outer membrane protein 16 (OEP16) (Reinbothe et al. 2004), which probably is derived from the Tim22 mitochondrial channel protein (Cavalier-Smith 2006).

The Reasons Behind Dinophysis *Kleptoplastid Longevity*

Although *Dinophysis* cells ingest only the inner plastid compartment from the cryptophyte plastids in *M. rubra*, these kleptoplastids are photosynthetically active for up to 2 months (Park et al. 2008). A similar longevity was reported for the cryptophyte kleptoplastids in *M. rubra* (Johnson et al. 2007), but in that case the long maintenance is explained by the presence of a transcriptionally active kleptokaryon (see section "Indirect *Myrionecta rubra*-Mediated Route of Cryptophyte Plastids into *Dinophysis* cells"). The cryptophyte nucleus and nucleomorph are completely absent from *Dinophysis* cells, meaning there must be alternative mechanisms for increasing their kleptoplastids' longevity.

Because Wisecaver and Hackett (2010) searched only a limited subset of the total *D. acuminata* transcriptome, it is possible that more plastid genes reside in the *Dinophysis* nucleus than estimated by these authors; however, it appears unlikely that the hundreds of additional plastid genes required for the long-term maintenance of *Dinophysis* kleptoplastids will be found. An alternative explanation of kleptoplastid longevity, however, could be an exceptional stability of proteins that reside in them. Such prolonged stabilities could be suggested for kleptoplastids in the foraminiferan genus *Elphidium* (Correira and Lee 2002a) and the sea slug *Elysia chlorotica* (Rumpho et al. 2011; Wägele and Martin 2013). In the case of *E. chlorotica*, Green et al. (2005) showed that kleptoplastids derived from the xanthophycean alga *Vaucheria litorea* are highly stable, with up to 30% of them intact after 14 days of incubation in light. Even greater longevity was observed for diatom plastids sequestered by the deep-sea foraminiferan *Nonionella stella* (Grzymski et al. 2002), which display very low turnover and can be retained up to 12 months.

Nature of Plastid Replacement in Dinophysis *Species*

The few plastid-related sequences found in *D. acuminata* are involved in distinct functions (Wisecaver and Hackett 2010). They participate in photosystem stabilization and protection (PsbM, PsbU, LI818), electron transfer (ferredoxin), and

metabolite transport (TPT). Only one of these proteins, PsbM, has a cryptophyte origin (Wisecaver and Hackett 2010); LI818 and PsbU appear to be derived from either haptophytes or fucoxanthin dinoflagellates, whereas ferredoxin and TPT group with peridinin dinoflagellate sequences in phylogenetic trees (Wisecaver and Hackett 2010). Although Wisecaver and Hackett (2010) considered the possibility that the *ferredoxin* and *TPT* genes were horizontally transferred from peridinin plastid-containing dinoflagellates to the *Dinophysis* lineage, they also could have been inherited vertically. *Dinophysis* species emerged from within the tree of peridinin dinoflagellates (Zhang et al. 2008; Qiu et al. 2011), indicating their ancestor contained such a plastid. At present, it is unclear whether any photosynthetic or nonphotosynthetic *Dinophysis* species still maintain a reduced peridinin plastid.

Dinoflagellates Containing Diatom-Derived Photosynthetic Bodies: Permanent Endosymbionts or True Cell Organelles?

General Characteristics of Dinotom Dinoflagellates

A number of dinoflagellate species are known to harbor an endosymbiotic alga of stramenopile origin (Fig. 3B; Table 1) (see, for example, Tomas and Cox 1973; Jeffrey and Vesk 1976; Tamura et al. 2005; Takano et al. 2008). Various lines of evidence presented in this section demonstrate that the endosymbionts are diatoms, which led use of the term "dinotom" after the two organisms involved (*dino*flagellate and dia*tom*) (Imanian et al. 2010). Dinotoms belong to at least seven different genera, including *Dinothrix*, *Durinskia*, *Galeidinium*, *Gymnodinium*, *Kryptoperidinium*, *Peridiniopsis*, and *Peridinium* (Table 2). These genera formerly were classified within the order Peridiniales but recently were moved to the Dinotrichales (Gómez 2012b). The current taxonomy of dinoflagellates with diatom-derived endosymbionts should be considered tentative and is likely to change with more extensive sampling, research, and discoveries of new species. For example, *Kryptoperidinium foliaceum* formerly was classified within different genera as *Glenodinium foliaceum* and *Peridinium foliaceum* (Gómez 2012b).

Dinotoms display great variation in morphology (including both athecate forms and thecate species with different plate configurations), habitats (freshwater to marine), and lifestyles (including planktonic and both motile and sessile benthic forms) (Horiguchi 2004, 2006). Some, such as *K. foliaceum*, *Peridiniopsis* spp., *Peridinium quinquecorne*, can form harmful blooms (Kempton et al. 2002; Garate-Lizarraga and Muneton-Gomez 2008; Zhang et al. 2011). Although production of toxins has not been reported, these blooms can cause noxious odors and produce fish kills by depleting the water of dissolved oxygen.

Table 2 List of dinotom endosymbionts/plastids discussed in this chapter

Dinotom genus	Dinotom species	Reference
Kryptoperidinium	*K. foliaceum*	Inagaki et al. (2000), Imanian and Keeling (2007), Figueroa et al. (2009), Imanian et al. (2010)
Durinskia	*D. baltica*	Inagaki et al. (2000), Imanian and Keeling (2007), Imanian et al. (2010)
	D. capensis	Pienaar et al. (2007)
Peridinium	*P. quinquecorne*	Horiguchi and Takano (2006)
Gymnodinium	*G. quadrilobatum*	Horiguchi and Pienaar (1994a)
Galeidiniium	*G. rugatum*	Tamura et al. (2005)
Dinothrix	*D. paradoxa*	Horiguchi (2006)
Peridiniopsis	*P.* cf. *kevei*	Takano et al. (2008)
	P. penardii	Takano et al. (2008)
	P. niei	Zhang et al. (2011)

Interestingly, the nonphotosynthetic species *Crypthecodinium cohnii* also is grouped in the order Dinotrichales, although in a separate family (Gómez 2012b). A close relationship between this species and *K. foliaceum* is supported by recent phylogenetic studies (Hoppenrath and Leander 2010). A number of proteins in *C. cohnii* show phylogenetic affinities to cyanobacterial or other algal homologs and some even contain intact plastid-targeting peptides typical of peridinin dinoflagellates (Sanchez-Puerta et al. 2007a). This suggests that *C. cohnii* may still contain an unrecognized, reduced plastid.

Ultrastructure and Origin of the Diatom-Derived Endosymbionts

Dinotom endosymbionts retain almost all organelles and intracellular structures of their free-living algal ancestors, including plastids, a nucleus, mitochondria, endoplasmic reticulum (ER), Golgi apparatus, and cytosol with ribosomes; the only exception is centrioles (Fig. 3B) (Tomas and Cox 1973; Jeffrey and Vesk 1976; Horiguchi and Pienaar 1994a; Tamura et al. 2005; Horiguchi and Takano 2006; Pienaar et al. 2007; Takano et al. 2008). Each dinotom host cell harbors a single endosymbiont, which occupies most of its volume. The endosymbiont can be quite expansive, containing 10–20 plastids in its cytosol (Dodge 1983; Tamura et al. 2005). Dinotom endosymbionts are surrounded by a single membrane separating them from the host cytosol (see, for example, Tomas and Cox 1973; Jeffrey and Vesk 1976; Tamura et al. 2005; Takano et al. 2008); its evolutionary origin, phagosomal or endosymbiont plasmalemma, remains unclear (for a discussion, see Eschbach et al. 1990; Schnepf and Elbrächter 1999).

From initial observations, ultrastructural and biochemical features of dinotom plastids clearly indicated that they were derived from a stramenopile alga (Tomas and Cox 1973; Jeffrey and Vesk 1976; Withers et al. 1977); they are surrounded by

four membranes within the endosymbiont, the outermost studded with ribosomes and connected with the nuclear envelope. They are devoid of a nucleomorph. Their thylakoids are stacked in threes and are surrounded by a girdle lamella, and the major photosynthetic pigments are chlorophylls *a* and *c* and the carotenoid fucoxanthin. Based on all these characteristics, it was suggested that dinotom endosymbionts probably evolved either from a chrysophyte or a diatom (Tomas and Cox 1973; Jeffrey and Vesk 1976; Withers et al. 1977; Dodge 1983; Kite and Dodge 1985). Their specific evolutionary origin finally was clarified by phylogenetic analyses of genes residing within their nuclei and plastids, which clearly demonstrate their diatom ancestry (see, for example, Chesnick et al. 1996, 1997; Tamura et al. 2005; Takano et al. 2008).

One of the most interesting features of dinotoms is that both the endosymbiont and dinoflagellate host mitochondria have been left relatively unchanged by the symbiosis (Imanian and Keeling 2007; Imanian et al. 2012). Sequences from the host mitochondria in *Durinskia baltica* and *K. foliaceum* revealed the presence of typical dinoflagellate features (Imanian et al. 2012). For example, both appear to contain the same three protein genes encoding cytochrome b (*cob*), two cytochrome oxidase subunits (*cox1* and *cox3*) and several fragments of the large subunit of ribosomal RNA (*LSU rRNA*). Moreover, the host mitochondrial genomes are expanded dramatically in size through recombination and amplification of numerous pseudogenes. The only substantive difference found in either was a unique 150 bp insertion in the *cob* gene from *D. baltica* (Imanian et al. 2012). Likewise, the endosymbionts' mitochondrial genomes retain features typical of diatoms. Neither has been reduced in size and both are highly conserved in gene content and gene order relative to their free-living diatom counterparts (Imanian et al. 2012). Perhaps most significant, both the host and endosymbiont mitochondria remain metabolically active and exhibit overlapping functions (Imanian and Keeling 2007). This was demonstrated by expression analyses of various endosymbiont mitochondrial genes in *K. foliaceum*, including *cox1*, *cox2*, *cox3*, *cob*, and *LSU rRNA*, as well as *cox1* and *cob* from the host cell's mitochondrion (Imanian and Keeling 2007).

Despite the maintenance of nearly all intracellular structures, dinotom endosymbionts have lost some important features, such as a silicious cell wall (or frustule), motility, and a visible mitotic stage (Tomas and Cox 1973; Jeffrey and Vesk 1976; Tippit and Pickett-Heaps 1976; Chesnick and Cox 1989; Tamura et al. 2005; Takano et al. 2008; Figueroa et al. 2009). Consequently, they bear no obvious resemblance to their free-living diatom relatives. The endosymbiont's division is synchronized with the host cell's and proceeds by constriction (Tippit and Pickett-Heaps 1976; Chesnick and Cox 1989; Figueroa et al. 2009); however, there is no evidence of condensation of chromosomes or tight control over their segregation into daughter nuclei. Moreover, it was demonstrated that dinotom endosymbionts do not synthesize complex polysaccharides for energy storage, instead relying on starch stored by the dinoflagellate host (Dodge 1983). All of these observations demonstrate a deep integration between the endosymbiont and its host cell; however, it remains debatable whether the diatom endosymbionts

should be considered true cell organelles (see section "Putative Protein Import into Dinotom Endosymbionts").

Complex Evolutionary History of Dinotom Endosymbionts

Monophyletic Versus Polyphyletic Origins of the Diatom-Derived Endosymbionts

All dinotoms share some intracellular features, such as a peculiar eyespot (Horiguchi and Pienaar 1994a), but they are quite diverse in their morphologies, lifecycles, and habitats (for a review, see Horiguchi 2006). Based on these differences, it is reasonable to hypothesize that they represent a polyphyletic assemblage and acquired diatom endosymbionts independently on multiple occasions (Fig. 5). Although most phylogenetic analyses indicate monophyly of dinotom host cells (Inagaki et al. 2000; Tamura et al. 2005; Horiguchi and Takano 2006; Pienaar et al. 2007; Takano et al. 2008), some studies cast doubt on this conclusion (Morris et al. 1993; Zhang et al. 2011). Moreover, phylogenies based on endosymbiont genes clearly demonstrate multiple origins of the diatom endosymbionts, which were further complicated by replacement events (Horiguchi and Takano 2006; Takano et al. 2008; Zhang et al. 2011). Thus, the evolutionary history of endosymbionts residing in dinotoms appears to be much more complex (Fig. 5) than initially supposed.

Establishment of the Pennate Diatom Endosymbiont

Most phylogenetic trees of nuclear genes indicate that dinotoms such as *K. foliaceum*, *Galeidinium rugatum*, *Durinskia capensis*, and *D. baltica* are monophyletic (Inagaki et al. 2000; Tamura et al. 2005; Pienaar et al. 2007). This is consistent with a scenario in which a common ancestor of the clade engulfed a diatom and established it as a permanent endosymbiont (Fig. 5). It is evident that the alga involved was a pennate diatom from the family Bacillariaceae, which includes genera such as *Nitzschia*, *Pseudonitzschia*, *Bacillaria*, and *Cylindrotheca* (Tamura et al. 2005; Pienaar et al. 2007; Chesnick et al. 1996, 1997). Although it usually is assumed that the pennate endosymbiont was *Nitzschia*-like, Tamura et al. (2005) showed that it also bears a resemblance to *Cylindrotheca* sp.

Most members of the Bacillariaceae are benthic algae, living on shallow marine sediments, whereas dinotoms are mainly planktonic forms (for a detailed discussion, see Chesnick et al. 1997). Thus, if a benthic pennate diatom was the ancestor of dinotom endosymbionts, the question arises as to how planktonic dinoflagellates encountered their future endosymbionts in the first place? Chesnick et al. (1997) proposed two possible scenarios. In one, the production of hypnozygotes (benthic

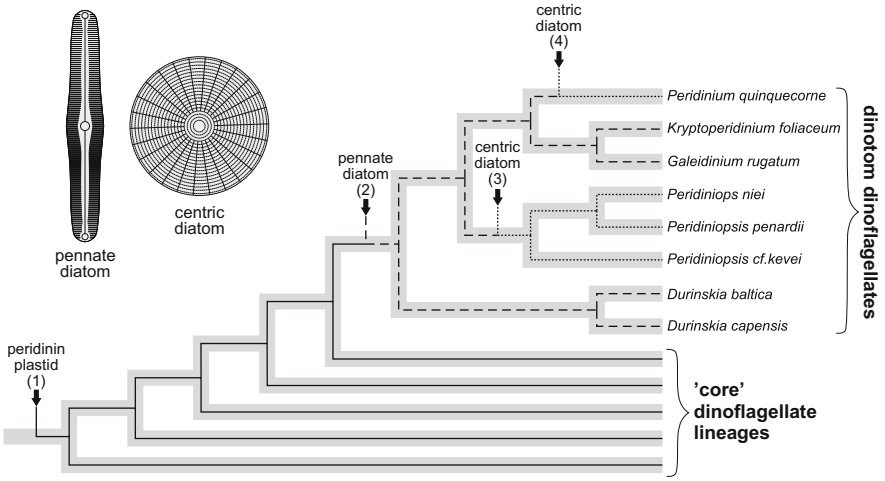


Fig. 5 Complex evolutionary pathway of diatom endosymbionts in dinotom dinoflagellates. It is very likely that the ancestor of all these dinoflagellates had a peridinin plastid based on its very early acquisition in the evolution of "core" dinoflagellates (1). Additional support for this hypothesis is the presence of a peculiar type B eyespot in dinotom cells, which probably evolved from the peridinin plastid after the acquisition of a new pennate diatom endosymbiont related to genera like *Nitzschia*, *Bacillaria* or *Cylindrotheca* (2). The descendent of this endosymbiont is present in *Durinskia* spp., *Kryptoperidinium foliaceum*, and *Galeidinium rugatum*; however, it was replaced, independently, in the *Peridiniopsis* lineage (3) and in *Peridinium quinquecorne* (4) by centric diatoms related to *Thalassiosira*/*Skeletonema* and *Chaetoceros*, respectively. Note that these presumed evolutionary relationships among dinotom species are based on poorly resolved trees (Takano et al. 2008; Zhang et al. 2011); thus, it is possible that acquisition of a centric diatom endosymbiont by *Peridiniopsis* dinoflagellates involved direct replacement of the peridinin plastid, rather than proceeding through a pennate diatom intermediate

sexual cysts formed after gamete fusion) could have facilitated encounters between benthic diatoms and an ancestor of dinotom host cells. Alternatively, the host cell ancestor of dinotoms could have been a bottom-dwelling dinoflagellate like, for example, *Gymnodinium quadrilobatum* (Horiguchi and Pienaar 1994a).

Independent Replacements of the Pennate Endosymbiont by Centric Diatoms

Available data indicate that the original pennate endosymbiont was replaced by distinct centric diatoms, independently, in two separate lineages (Fig. 5). The first of these lineages, represented by *P. quinquecorne*, groups with pennate diatom endosymbiont-containing dinotoms in phylogenetic analysis of host cell nuclear small subunit (SSU) rDNA (Fig. 5) (Horiguchi and Takano 2006). Phylogenies based on *P. quinquecorne* endosymbiont genes, however, show it to be derived from a centric rather than pennate diatom, and probably one related to the genus *Chaetoceros* (Horiguchi and Takano 2006). Because *P. quinquecorne* emerges

from within the dinotom clade in nuclear SSU rDNA trees, it is reasonable to postulate that its ancestor initially harbored the same pennate endosymbiont present in most other dinotoms, which later was replaced by a *Chaetoceros*-like centric diatom (Horiguchi and Takano 2006).

A second case of endosymbiont replacement appears to have occurred in *Peridiniopsis* cf. *kevei* and *P. penardii*, which, in contrast to other dinotom species found in marine environments, were isolated from freshwater habitats in Japan (Fig. 5) (Takano et al. 2008). On phylogenetic trees of nuclear genes, these species also group with typical "pennate" dinotoms, whereas phylogenies of endosymbiont genes indicate a close relationship to centric diatoms like *Thalassiosira* or *Skeletonema* (Takano et al. 2008). These data suggest that, as in *P. quinquecorne*, a common ancestor of *P.* cf. *kevei* and *P. penardii* originally had the pennate endosymbiont that was later replaced by the centric diatom (Takano et al. 2008).

The results of initial studies (Takano et al. 2008) were complicated by subsequent and expanded investigations of *Peridiniopsis* species by Zhang et al. (2011), which included *P. penardii* var. *robusta*, as well as two morphospecies of *P. penardii* and a new species, *P. niei*, from China (Fig. 5). Phylogenetic analyses of endosymbiont nuclear genes not only confirmed a centric thalassiosiroid diatom origin but also suggested that the genus *Discostella* is a closer relative of these endosymbionts than *Thalassiosira* or *Skeletonema* (Zhang et al. 2011). The freshwater *Peridiniopsis* dinotom species formed a monophyletic clade in all trees based on host cell nuclear genes (Zhang et al. 2011); however, the relationship of this clade to other diatom endosymbiont-containing dinoflagellates was uncertain. On the SSU rDNA tree all these dinotoms, along with *Peridiniopsis* species, formed a single assemblage; however, in a tree of ITS rDNA sequences *Peridiniopsis* represented a separate clade, thereby suggesting an independent acquisition of the thalassiosiroid endosymbiont from the pennate endosymbiont (Zhang et al. 2011).

Genomics of Diatom Endosymbiont Plastids

A diatom ancestry of dinotom endosymbionts is further supported by complete plastid genome sequences from two closely related dinotoms, *D. baltica* and *K. foliaceum* (Imanian et al. 2010). Numerous features, including general structure, ordered gene blocks, and gene content closely resemble those of the free-living pennate diatoms. The sizes of the circular genomes of *D. baltica* (116,470 bp) and *K. foliaceum* (140,426 bp) are in the range of those present in the diatoms *Phaeodactylum tricornutum* (117,369 bp) (Oudot-Le Secq et al. 2007) and *Fistulifera* sp. (134,918 bp) (Tanaka et al. 2011). All the four dinotoms and diatoms have comparable low overall G+C contents (32–33 %). Even more telling, however, is that G+C percentages show correlated differences across classes of genes among the diatom endosymbiont and free-living diatom plastids. For example, tRNA genes all have higher than average G+C content (52–54 %), whereas

intergenic regions have much lower G+C fractions (16–26 %) (Imanian et al. 2010). There are comparable numbers of protein-encoding genes (*D. baltica*, 127; *K. foliaceum*, 128; *P. tricornutum*, 130; and *Fistulifera* sp., 132), identical numbers from all classes of RNA genes (3 rRNAs, 27 tRNAs, and 2 other stable RNAs), and the same four overlapping pairs of genes in all the diatom endosymbiont and free-living diatom plastid genomes examined to date; in addition, all are completely devoid of introns (Imanian et al. 2010; Tanaka et al. 2011).

Overall structure and gene order are highly conserved between the plastid genomes from *D. baltica* and *P. tricornutum*, with only a few inversions, translocations, and gene deletions distinguishing the two (Imanian et al. 2010). Average distances between genes also are quite similar, 94.3 bp in *D. baltica* versus 88.4 bp in *P. tricornutum*. Interestingly, the average intergenic distance is much larger (246.7 bp) and there are far more genomic rearrangements in the diatom plastid genome sequenced from *K. foliaceum* (Imanian et al. 2010). Both of these differences appear to be related to the presence of nine large insertions with significant similarity to the pCf1 and pCf2 plasmids in the pennate diatom *Cylindrotheca fusiformis* (Imanian et al. 2010) that appears to be closely related to the diatom endosymbionts in dinotoms (see section "Establishment of the Pennate Diatom Endosymbiont"). This includes apparently functional homologs of serine recombinases *serC1* and *serC2*, as well as two smaller truncated copies of *serC1*. Moreover, based on sequence similarities, it appears that the two genes in the *K. foliaceum* plastid genome are paralogs, each related more closely to a respective ortholog on one of the two *Cylindrotheca* plasmids (Imanian et al. 2010). In addition, several dispersed stretches of plastome sequence in *K. foliaceum* show strong similarity to noncoding regions of the *Cylindrotheca* plasmids. Although probably unrelated, the *K. foliaceum* plastid genome also contains a site-specific tyrosine recombinase gene with strong sequence similarity to the *tyrC* gene from the raphidophyte alga *Heterosigma akashiwo* (Imanian et al. 2010). All of these insertions have resulted in more substantial genomic changes in the *K. foliaceum* plastid since it diverged from those in *D. baltica* and free-living diatom relatives.

A homolog of the *serC2* gene and other sequences related to the *Cylindrotheca* plasmids also were found to be incorporated into the plastid genome of the diatom *Fistulifera* sp. (Tanaka et al. 2011). Presumably, as in *K. foliaceum*, these plasmid insertions have contributed to an increased overall size of the *Fistulifera* plastid genome relative to other free-living diatoms and to a greater average intergenic distance (179.5 bp). When and how these pCf1 and pCf2 sequences were acquired by *Fistulifera* and *Kryptoperidinium* plastid genomes is unclear. Interestingly, however, the *serC2* gene from *Fistulifera* shows greater sequence similarity to its homolog in *K. foliaceum* than to homologs from either plasmid in *C. fusiformis* (Tanaka et al. 2011). This tends to suggest that unintegrated pCf1 and pCf2 plasmids were present in the common ancestor of *Fistulifera* and dinotom endosymbionts, which later were integrated independently into the plastid genomes in *Fistulifera* and *K. foliaceum*, but lost without integration from *D. baltica*.

The great similarities between plastid genomes from dinotoms and free-living diatoms are particularly striking given the relative evolutionary distances currently available for examination. Even fewer genomic changes are to be expected when free-living pennate diatoms that are more closely related to the ancestral dinotom endosymbionts are compared. The relative conservation of dinotom endosymbionts and their plastid genomes likely reflects a quite recent timing of the endosymbiosis, which has provided little time for substantive gene transfer, reorganization, or reduction. This stands in sharp contrast with genomes of haptophyte-derived tertiary plastids found in the dinoflagellate *Karlodinium veneficum* (Gabrielsen et al. 2011), which underwent massive rearrangements and gene losses compared to representatives of their potential donor taxa (see section "Evidence for Haptophyte Origin of *Karenia* and *Karlodinium* Plastids").

Putative Protein Import into Dinotom Endosymbionts

The presence of all cell organelles and nearly all subcellular structures in the endosymbiotic diatom suggests that they could be maintained without importing any proteins encoded in the dinoflagellate host nuclear genome. However, the disorganized division of the endosymbiont nucleus (Tippit and Pickett-Heaps 1976; Chesnick and Cox 1989; Figueroa et al. 2009), which results in variable sizes of descendant nuclei (with possible loss of important genes), suggests that some gene transfer from the endosymbiont to the host nucleus could be occurring (for a discussion, see Kite et al. 1988; Figueroa et al. 2009; Nowack and Melkonian 2010). Because the diatom endosymbionts still contain plastids, a nucleus and mitochondria, genes from any of them could move into the host cell's nuclear genome. Such transfers of cytosolic and nuclear genes have been confirmed in both the cryptophyte and chlorarachniophyte plastid endosymbioses (Curtis et al. 2012), despite the retention of a nucleomorph in both cases (Gilson et al. 2006; Lane et al. 2007; see also Tanifuji and Archibald 2013). Consequently, in addition to plastid-specific proteins, nuclear-encoded proteins also are imported into the periplastidal compartment and the nucleomorph (Gould et al. 2006; Hirakawa et al. 2010, 2011). In the case of dinotoms, proteins encoded in the dinoflagellate nucleus could be imported into the diatom's plastids, cytosol, nucleus, and mitochondria (Fig. 4B).

Regardless of their final destinations, these hypothetical endosymbiont-targeted proteins would have to cross the single endosymbiont-delimiting membrane (Fig. 4B). Thus, an important question arises: what transport system might operate in that membrane? To answer this issue requires consideration of the likely origin of the membrane surrounding dinotom endosymbionts. Schnepf and Elbrächter (1999) argued that it is derived from the host phagosomal membrane, implying the diatoms were engulfed via myzocytosis, during which the endosymbiont plasmalemma was lost. In accordance with this hypothesis, some signal peptide-carrying host proteins destined to the endosymbiont cytosol or nucleus could be delivered to the

surrounding endosymbiont membrane in vesicles derived from the host endomembrane system and then released into the endosymbiont cytosol (Fig. 4B). In the case of typical diatom plastid-targeted proteins equipped with bipartite presequences composed of a signal peptide followed by a transit peptide (Gruber et al. 2007), their signal peptides should remain uncut in the host ER (e.g., as a result of incompatibility of their cleavage motifs with host signal peptidases) to make them useful in the next translocation step across the plastid ER membrane (Fig. 4B). It is possible that the membrane contains a Sec61 channel associated with receptor proteins that are homologous to the yeast and mammalian Sec62/63 complex (Park and Rapoport 2012). This would permit recognition and translocation of imported plastid proteins with previously uncleaved signal peptides into the lumen of plastid ER. After cleavage of these peptides by compatible signal peptidases, the proteins would traverse the subsequent periplastidal membrane with the help of an ERAD-like translocon (or SELMA), and the two innermost membranes through Toc and Tic translocons (Fig. 4B) (Bullmann et al. 2010; Hempel et al. 2010; see also Grosche et al. 2013).

Freeze-fracture studies of the membrane surrounding the pennate diatom endosymbionts in *D. baltica* and *K. foliaceum*, however, indicated that the membrane represents the endosymbiont plasmalemma (Eschbach et al. 1990). This would suggest that diatoms were ingested via phagocytosis, after which the phagosomal membrane was lost, immediately and efficiently protecting the engulfed algae from being targeted for intracellular digestion. Considering this scenario, McEwan and Keeling (2004) suggested that the diatom's plasmalemma could have preserved its primary pinocytotic activity that was adapted to import proteins into the endosymbiont's organelles. Trafficking of dinoflagellate nucleus-encoded proteins into the diatom endosymbiont plastids could work as follows (Fig. 4B). First, specific pinocytotic vesicles with plastid proteins carrying only transit peptides would pinch off from the endosymbiont's surrounding membrane and fuse with the outermost plastid membrane, releasing the imported proteins into the lumen of plastid ER. Subsequently, these proteins would be translocated through SELMA and the Toc and Tic translocons into the plastid stroma (Bullmann et al. 2010; Hempel et al. 2010; Grosche et al. 2013). Because proteins imported into diatom plastids carry both signal and transit peptides (Gruber et al. 2007), the diatom endosymbiont plastid-targeted proteins would have to lose these peptides after transfer of their genes into the dinoflagellate host nuclear genome (Fig. 4B). Otherwise, the proteins would be cotranslationally translocated into the host endomembrane system, thereby preventing them from direct encounters with the surrounding endosymbiont membrane.

Although a pinocytotic pathway can explain import of plastid proteins, it does not apply to endosymbiont cytosolic and nuclear proteins. Thus, it is reasonable to postulate that the membrane surrounding the diatom endosymbionts contains a protein-conducting import pore for distinct kinds of endosymbiont-targeted proteins (Fig. 4B). Such channels could be derived from the host or the endosymbiont. The hypothetical existence of a protein channel-based transport system in the surrounding endosymbiont membrane would solve the problems with import of

distinct protein types into the diatom endosymbionts; however, it is clear that protein import into the diatom endosymbionts needs further investigations, especially given the generally unreduced state of the endosymbiont nucleus.

Peculiar Eyespot, a Remnant of the Peridinin Plastid?

Dinoflagellates have five types of eyespots (Horiguchi and Pienaar 1994b; Kreimer 1999). One of them, termed type B eyespot (Dodge 1984), is characteristic of dinotoms and is found in their cytosol (see, for example, Horiguchi and Pienaar 1994a; Tamura et al. 2005; Takano et al. 2008). The eyespot is composed of several layers of red-pigmented lipid globules that are enclosed by three membranes (Dodge and Crawford 1969; Dodge 1984). A very similar type C eyespot is present in some peridinin plastids (Dodge 1984; Kreimer 1999). Considering these similarities, Dodge (1984) and Horiguchi and Pienaar (1994a) suggested that the type B eyespot is the reduced remnant of the peridinin plastid. According to this evolutionary scenario, the ancestor of dinotoms, containing peridinin plastids with a type C eyespot, engulfed a diatom and established it as a permanent photosynthetic endosymbiont (Horiguchi and Pienaar 1994a). Because its primary metabolic contribution as a source of fixed carbon no longer was required, the peridinin plastid was free to transform, finally evolving into the type B eyespot.

Based on the common presence of type B eyespots, Horiguchi and Pienaar (1994a) argued for a single origin of dinotom dinoflagellates. They reasoned that multiple origins of dinotoms meant that the series of conditions and complex evolutionary events leading to these eyespots had to have occurred independently in each lineage. These would have included (1) the presence of a type C eyespot in each host cell, (2) ingestion and retention of a diatom alga, (3) establishment of a permanent endosymbiosis with the alga, (4) conversion of the peridinin plastid into a type B eyespot, and (5) multiplication of the initial layer of lipid globules (Horiguchi and Pienaar 1994a). Therefore, the type B eyespot should be considered a clear synapomorphic character uniting dinotoms as a natural, monophyletic group. These arguments of Horiguchi and Pienaar (1994a) are compatible with the molecular phylogenetic results, which suggest that dinotom host cells are monophyletic (see, for example, Horiguchi and Takano 2006; Pienaar et al. 2007; Takano et al. 2008), as well as replacements of diatom endosymbionts in *P. quinquecorne* and *Peridiniopsis* species described above.

True Plastids of Haptophyte Origin in Fucoxanthin Dinoflagellates

General Characteristics of Kareniaceae *Dinoflagellates*

Three dinoflagellate genera, *Karenia*, *Karlodinium*, and *Takayama*, comprise a distinct evolutionary lineage, for which a new family was suggested, the *Kareniaceae* (Bergholtz et al. 2006). The phylogenetic position of the *Kareniaceae* relative to other dinoflagellates, however, has not been pinned down. Phylogenetic analyses based on mitochondrial cytochrome *b* and cytochrome *c* oxidase placed them within "core" dinoflagellates (Zhang et al. 2007b), whereas those based on rDNA and Hsp90 (Hoppenrath and Leander 2010) as well as combined sets of rDNA, mitochondrial, and nuclear protein genes (Orr et al. 2012), recovered them as one of the earliest diverging branches of "core" dinoflagellates. Traditionally, the *Kareniaceae* were classified in the Gymnodiniales but recently were moved to the *Brachidiniaceae* within the Brachidiniales (Gómez 2012b).

The two best-known representatives of this family, *Karenia brevis* and *Karlodinium veneficum*, are the primary focus of this discussion. They are unarmored, mixotrophic species (Adolf et al. 2006) with either straight or "S"-shaped apical grooves (Daugbjerg et al. 2000; Bergholtz et al. 2006), and live in marine (and estuarine, in the case of *K. veneficum*) environments (Brand et al. 2012). They can form large blooms, usually in coastal waters, that produce a variety of compounds toxic to fish, invertebrates, and other marine organisms (Hackett et al. 2004a; Brand et al. 2012; Zhou et al. 2011).

Karlodinium toxins are linear, amphipathic polyketides, called karlotoxins, and glycoglycerolipids (e.g., digalactosyldiacylglycerol) with hemolytic, cytotoxic, and ichthyotoxic activities (Bachvaroff et al. 2008; Zhou et al. 2011). *Karenia* is best known for production of brevetoxin, but species in this genus also can produce gymnodimine, gymnocins, and a variety of sterols and polyunsaturated fatty acids that have allelopathic effects on other phytoplankton (Brand et al. 2012). Brevetoxins are cyclic polyethers that can impact many neurological processes and are responsible for neurotoxic shellfish poisoning (NSP); they also can cause respiratory distress when aerosolized (Watkins et al. 2008).

Evidence for Haptophyte Origin of Karenia *and* Karlodinium *Plastids*

A haptophyte-derived plastid with unique pigmentation is present in *Karenia*, *Karlodinium*, and *Takayama* (Fig. 3C; Table 1) (for a review, see Schnepf and Elbrächter 1999). The plastid contains chlorophyll *c* along with the carotenoid fucoxanthin and its derivatives, such as 19′-hexanoylofucoxanthin, 19′-butanoyloxyfucoxanthin,

and 19-hexanoyloxyparacentrone 3-acetate (gyroxanthindiester) (Hansen et al. 2000a; Bjørnland et al. 2003; de Salas et al. 2003). These plastids are quite different from peridinin plastids found in most dinoflagellates, which contain the carotenoid peridinin rather than fucoxanthin and its derivatives (Schnepf and Elbrächter 1999); consequently, they are called "fucoxanthin plastids." This unusual pigmentation, particularly the fucoxanthin derivatives, is characteristic of prymnesiophyte or haptophyte algae like *Chrysochromulina* (Zapata et al. 2004). Additional features of plastid ultrastructure (no girdle lamella, thylakoids stacked in threes, and internal pyrenoids) and multiple phylogenetic studies of plastid genes, all indicate that plastids in the *Kareniaceae* are derived from an endosymbiosis involving a haptophyte alga (Fig. 3C) (Schnepf and Elbrächter 1999; Tengs et al. 2000; Yoon et al. 2002; Ishida and Green 2002; Takishita et al. 2004; Patron et al. 2006; Shalchian-Tabrizi et al. 2006).

Initial phylogenies based on the plastid genes *psa*A, *psb*A, and *rbc*L suggested a common haptophyte origin of the fucoxanthin and peridinin plastids (Yoon et al. 2002); however, it was pointed out that convergent codon usage in *psbA* in haptophytes and dinoflagellates could have led to their artificial clustering in trees based on these three genes (Inagaki et al. 2004; Inagaki and Roger 2006). Phylogenies correcting for this problem, and those based on amino acid sequences, decreased the significance for the haptophyte-peridinin plastid grouping, but provided only very weak support for an alternative relationship between peridinin and stramenopile plastids (Inagaki et al. 2004; Inagaki and Roger 2006). More gene-rich plastid phylogenies provided stronger support for grouping peridinin plastids with those from stramenopiles (Yoon et al. 2005; Wang et al. 2008) or apicomplexans and chromerids (Janouskovec et al. 2010). The application of more appropriate phylogenetic methods (e.g., employing covarion models) based on amino acid and recoded nucleotide sequences, combined with removal of the controversial *psbA* and other fast-evolving or compositionally biased sequences, recovered strong support for a peridinin-haptophyte plastid clade (Sanchez-Puerta et al. 2007b; Shalchian-Tabrizi et al. 2006; Bachvaroff et al. 2005). Additional taxon sampling of haptophyte plastids indicated that the peridinin and fucoxanthin plastids were acquired independently from two respective haptophyte lineages, the *Pavlovophyceae* and *Prymnesiophyceae* (Shalchian-Tabrizi et al. 2006).

A haptophyte origin of fucoxanthin plastids also is supported by the discovery of many haptophyte genes in the *Karenia* and *Karlodinium* nuclear genomes (Yoon et al. 2005; Nosenko et al. 2006; Patron et al. 2006), which are consistent with a process of endosymbiotic gene transfer. Finally, this evolutionary scenario received strong confirmation from recent comparative genomics of the plastid genomes from *K. veneficum* and the haptophyte *Emiliania huxleyi* (Gabrielsen et al. 2011).

Karlodinium *Plastid Genome and Its Transformation*

Studies of the *K. veneficum* plastid genome not only supported its close relationship with the haptophyte group *Prymnesiophyceae* (Gabrielsen et al. 2011) but also revealed major changes with respect to the *E. huxleyi* plastid genome (Sanchez Puerta et al. 2005), undoubtedly resulting from transformations associated with the tertiary endosymbiosis in dinoflagellates. The two genomes are very close in G+C content (~27 %) and their rRNA and tRNA complements have not changed substantially; however, the *Karlodinium* plastid genome encodes only 70 of the 119 protein encoding genes annotated in *Emiliania*. Despite this substantial reduction in gene number, the *K. veneficum* plastid genome is actually larger (142,981 versus 105,309 bp) because of expanded intergenic regions (Gabrielsen et al. 2011). These regions constitute 51.3 % of the *Karlodinium* genome compared to only 13.2 % in *Emiliania*, with average lengths of 733 bp versus 97.6 bp, respectively.

The *K. veneficum* genome has been subjected to substantial rearrangements and deletions (Gabrielsen et al. 2011). For example, gene clusters like the *atpA* operon and ribosomal protein superoperon, which typically are conserved across red-algal derived plastid genomes, are incomplete, reduced, or fragmented. Perhaps most interesting, two rearrangements within the ribosomal superoperon appear to have occurred independently in both *K. veneficum* and in the red alga-derived plastid in the *Chromera velia* (Janouskovec et al. 2010). In both cases, *rpl31* was translocated and *rps13* and *rpl36* were inverted, resulting in the same gene cluster (*S13–L36–S11–L31*) that is not present in other known red alga-derived plastids. In addition, the rRNA operon has an unusual structure, in which 5S and 16S rDNAs were duplicated, but not 23S rDNA.

Many gene sequences in the *Karlodinium* plastid genome have changed considerably compared to their homologs in other plastid genomes (Tengs et al. 2000; Shalchian-Tabrizi et al. 2006; Gabrielsen et al. 2011). Ten genes have internal stop codons, five have frame shifts, five have lost substantive and conserved domains, and two have diverged significantly in overall sequence (Gabrielsen et al. 2011). A number of these genes could be subjected to RNA editing, as occurs in peridinin dinoflagellate plastids (Zauner et al. 2004; Wang and Morse 2006; Dang and Green 2009) and mitochondria (Lin et al. 2002); alternative explanations would be the use of a variant genetic code and/or the presence of numerous pseudogenes.

Gene loss, accelerated sequence evolution, and genome rearrangements are not unique to the *K. veneficum* plastid genome; they are found in peridinin dinoflagellate, apicomplexan, and chromerid plastid genomes as well (Zhang et al. 2000; Bachvaroff et al. 2006; Janouskovec et al. 2010). This suggests that similar evolutionary processes have occurred independently during plastid acquisition by different alveolate taxa (Gabrielsen et al. 2011). Moreover, presumably independent plastid genome reductions in *Karlodinium* and *Chromera* resulted in a largely overlapping gene content, including 12 genes for photosystem proteins that also are encoded in the even more reduced peridinin plastid genome (Gabrielsen et al. 2011). This could result from reductive processes observed more broadly in

plastid genome evolution (Stiller et al. 2002), which lead to convergence on a "core" set of plastid genes in multiple alveolate lineages (Gabrielsen et al. 2011).

The massive reorganization of haptophyte-derived plastid genomes clearly contrasts with only minor changes that have occurred in dinotom plastid genomes (see section "Genomics of Diatom Endosymbiont Plastids"). The different levels of change likely reflect different stages of endosymbiotic integration. In dinotoms, most of the secondary host cell is preserved (section "Ultrastructure and Origin of the Diatom-Derived Endosymbionts"), whereas the plastid is all that remains in *K. veneficum*. Consequently, the cellular environment in dinoflagellates, which seems to favor more substantial plastid genome reductions (Espelund et al. 2012), has had a far greater impact on the more derived fucoxanthin plastids. It is interesting that this greater integration occurred with a haptophyte but not a diatom replacement of the peridinin plastid. This could reflect a close phylogenetic affinity of the ancestors of peridinin and fucoxanthin plastids (both of probable haptophyte origin; see Shalchian-Tabrizi et al. 2006; Gabrielsen et al. 2011), permitting the latter to coopt mechanisms related to the former more easily (see section "Plastid Replacement in Fucoxanthin Dinoflagellates" below).

The dinoflagellate intracellular environment appears to have led to another remarkable example of convergence in the genomes of adopted plastids. In-depth sequencing of the *K. veneficum* plastid genome revealed extrachromosomal DNA containing genes and gene fragments, which also are present as complete sequences on the conventional plastid chromosome (Espelund et al. 2012). The most abundant are the chaperone Hsp70 (*dnaK*) and the RuBisCO large subunit (*rbcL*), mostly fragmented and with 3′ truncations, as well as two tRNAs (*trnE* and *trnM*). These fragments resemble what is found in peridinin plastid genomes, which are organized entirely into numerous, small chromosomes called minicircles (for reviews, see Howe et al. 2008; Lin 2011; Wisecaver and Hackett 2011). Interestingly, genes for photosystem proteins PsaB, PsaA, PsbB, and PsbD, which have been found on peridinin minicircles, are among the genes overrepresented in *K. veneficum* indicating their presence in the extrachromosomal DNA fraction. *Karlodinium* extrachromosomal plastid sequences differ from peridinin minicircles, however, in most of the genes present, their organization and in length. What has driven these independent reductions and fragmentations of plastid genomes is unclear; however, some host-driven processes, common and unique to dinoflagellates, appear to be the cause (Espelund et al. 2012).

How Many Membranes Surround Fucoxanthin Plastids: Two, Three, or Four?

The number of envelope membranes of fucoxanthin plastids remains unclear. For example, Tengs et al. (2000) reported a range from two to four; however, electron micrographs published by Steidinger et al. (1978), Kite and Dodge (1988), and

Hansen et al. (2000b) indicate that three envelope membranes are present, at least in *K. veneficum* and *K. brevis*. Additional support for this finding comes from the presequence structure of proteins imported into fucoxanthin plastids (Fig. 4C). Most carry classical bipartite N-terminal targeting signals composed of a signal peptide followed by a transit peptide (class II proteins) (Patron et al. 2006; Patron and Waller 2007; see also Grosche et al. 2013); however, there are also proteins with tripartite presequences (class I proteins). In addition to the signal and transit peptides, they contain a second hydrophobic domain located downstream of the transit peptide (Patron et al. 2006; Patron and Waller 2007; Grosche et al. 2013). The domain probably functions as a stop-transfer sequence, anchoring the protein in the ER membrane. Because the tripartite N-terminal targeting signals occur only in protists with three-membrane-bound plastids, such as peridinin dinoflagellates (Nassoury et al. 2003; Patron et al. 2005) and euglenoids (Durnford and Gray 2006; Sláviková et al. 2005), it is reasonable to postulate that the envelope of fucoxanthin plastids is composed of three membranes as well (Bodył and Moszczyński 2006).

Mode of Acquisition, Protein Targeting Mechanisms, and Membrane Homology of Fucoxanthin Plastids

Theoretically, fucoxanthin plastids could have been acquired via phagocytosis, but fucoxanthin dinoflagellates are equipped with a peduncle (Steidinger et al. 1998; de Salas et al. 2003), which clearly indicates that they practice myzocytotic feeding. Thus, it is most likely that fucoxanthin plastids were gained via myzocytosis (Fig. 3C), which also is suggested by the absence of haptophyte mitochondrial genes in the *Karlodinium micrum* nuclear genome (Danne et al. 2012). With phagocytosis, the whole haptophyte cell would be engulfed and, after digestion, some genes from degraded mitochondria would have migrated into the host nucleus. Because no such genes were detected in *K. micrum* (Danne et al. 2012), it is the more likely that the haptophyte cell engulfed was incomplete and devoid of mitochondria (Fig. 3C).

The presence of two plastid protein classes with distinct targeting signals in fucoxanthin dinoflagellates, as discussed above, suggests they are trafficked through different routes. Based on published work on plastid protein import in peridinin dinoflagellates and euglenoids (see, for example, Sulli et al. 1999; Nassoury et al. 2003; Sláviková et al. 2005), the following model could be proposed (Fig. 4C). Because both class I and class II proteins carry signal peptides, both undoubtedly are transported cotranslationally into the ER; however, their targeting pathways would diverge significantly during this initial step. Class II proteins would be translocated completely into the ER lumen and transported directly to the plastid in ER-derived vesicles, bypassing the Golgi apparatus. When these vesicles fuse with the outermost plastid membrane, the proteins would be released

into the space between the outermost and middle envelope membranes. From there they would cross the two innermost plastid membranes, probably through Toc and Tic translocons that recognize their transit peptides.

The targeting pathway of class I proteins seems to be more complex (Fig. 4C). Given the presence of a stop-transfer domain, it is reasonable to suggest that these proteins are cotranslationally anchored in the ER membrane, preventing their complete migration into the ER lumen. They would be delivered to the plastid in this membrane-bound state in transport vesicles trafficked through the Golgi apparatus. The transit peptide would protrude into the vesicle lumen, whereas the future mature protein would remain in the cytosol. Golgi-derived transport vesicles then would fuse with the outermost plastid membrane, bringing the class I proteins with them. Based on the model proposed for class I proteins in *Euglena* (Sulli et al. 1999), these proteins then would migrate laterally within the outermost envelope membrane until they encounter a Toc translocon in the middle membrane, where the transit peptide would be recognized and promote movement through the Toc channel and, subsequently, through the Tic translocon to the inside of the plastid.

After it was engulfed by myzocytosis, the algal ancestor of fucoxanthin plastids would have been surrounded by five membranes, the phagosomal membrane of the host and the four envelope membranes of the haptophyte plastid (Fig. 3C). Because extant fucoxanthin plastids appear to have a three-membrane envelope, two of the original five membranes must have been lost. The endomembrane system-mediated protein targeting to these plastids suggests that the phagosomal membrane was retained. Because proteins imported into fucoxanthin plastids carry transit peptides, it also is most parsimonious to assume that the two innermost membranes of the original haptophyte plastid, containing Toc and Tic translocons (Bullmann et al. 2010; see also Grosche et al. 2013) were kept. Consequently, the plastid ER and periplastidal membranes likely were the two lost.

Plastid Replacement in Fucoxanthin Dinoflagellates

Although the specific phylogenetic position of fucoxanthin dinoflagellates has been difficult to determine, they are recovered consistently among peridinin dinoflagellates (Fig. 1) (Zhang et al. 2007b; Orr et al. 2012). This suggests that a peridinin plastid was present in the ancestor of fucoxanthin dinoflagellates, and later was replaced by the haptophyte plastid (Fig. 6). According to the replacement model discussed previously in the section "Unusual Dinoflagellate Plastids and Their Evolutionary Pathways," the integration of a new plastid could have been facilitated by the presence of preexisting genes for proteins targeted to the original plastid; however, initial gene and genomic analyses did not support such a hypothesis in the case of fucoxanthin dinoflagellates. Ishida and Green (2002) demonstrated that the gene encoding oxygen-evolving enhancer 1 (PsbO) protein in *K. brevis* is of haptophyte origin. To explain this, they proposed that either the

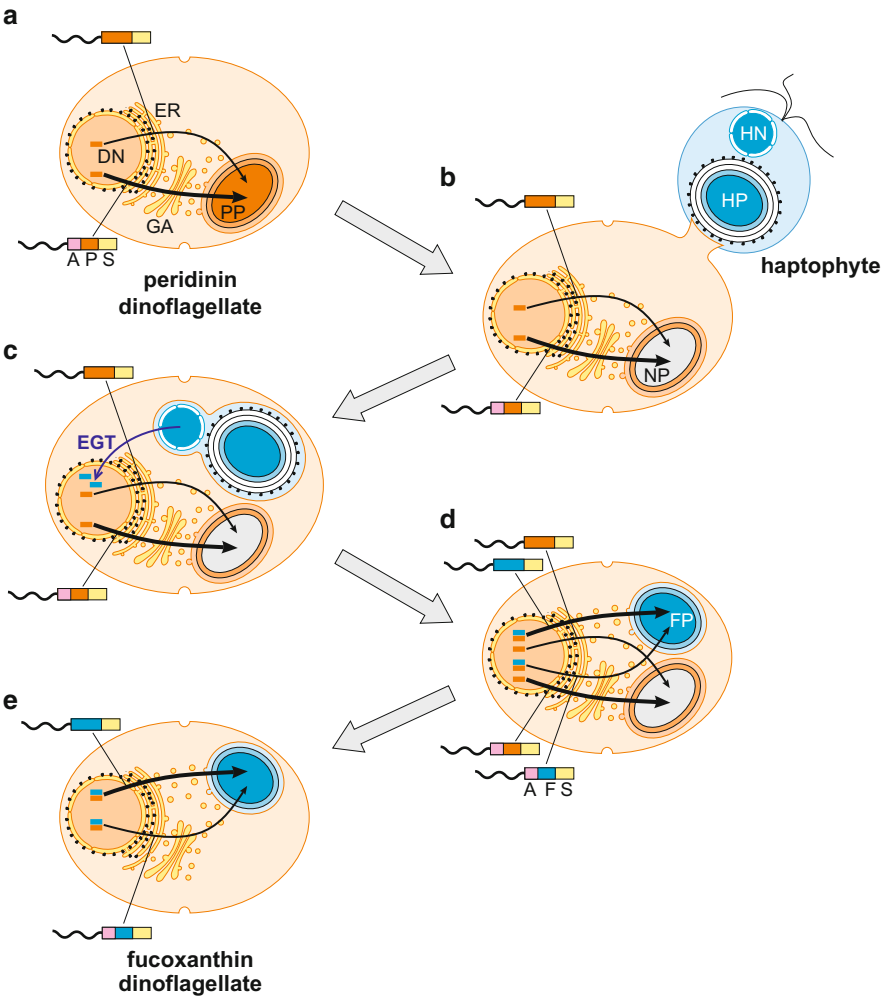


Fig. 6 Plastid replacement in fucoxanthin dinoflagellates. A: Phylogenetic analyses indicate that fucoxanthin dinoflagellates emerge from within peridinin plastid (PP) containing dinoflagellates. The majority of proteins encoded in the dinoflagellate nucleus (DN) and imported into peridinin plastids carry tripartite presequences comprising a signal peptide (S), transit peptide (P), and hydrophobic signal-anchor domain (A). These class I proteins probably are targeted to the peridinin plastid as integral membrane proteins via the endoplasmic reticulum (ER) and Golgi apparatus (GA). In contrast, class II plastid-targeted proteins lack the signal-anchor, and likely are targeted only through the ER. B: The peridinin plastid probably lost its photosynthetic activity, becoming a nonphotosynthetic plastid (NP) before acquisition of a haptophyte plastid, and was maintained for essential, nonphotosynthetic functions. The haptophyte plastid (HP) and nucleus (HN) likely were acquired via myzocytosis. C: Plastid-related genes from the haptophyte nucleus moved to the dinoflagellate nucleus via endosymbiotic gene transfer (EGT), finally resulting in disappearance of the haptophyte nucleus. D: The haptophyte plastid lost its two original outer membranes and was transformed into a three-membrane-bound fucoxanthin plastid (FP) capable of importing both peridinin and haptophyte plastid-related proteins. In contrast with proteins remaining from the earlier peridinin plastid, most haptophyte plastid proteins probably are directed to the new plastid via the ER-mediated pathway. The peridinin and fucoxanthin plastids could

peridinin plastid was lost before acquisition of the fucoxanthin plastid, or that the targeting systems of the two plastids were incompatible, meaning loss of the peridinin plastid also resulted in the loss of genes encoding its proteins. Similar results were obtained for a plastid-targeted isoform of GAPDH. Takishita et al. (2004) demonstrated that this *gapC1-fd* gene is of a haptophyte and not dinoflagellate origin. Also, more extensive studies of ESTs found no peridinin plastid-related genes in *K. brevis* (Yoon et al. 2005). Those missing included the gene for proteobacterial form II of RuBisCO that is characteristic of peridinin dinoflagellates (Whitney et al. 1995). Based on these findings, Yoon et al. (2005) concluded that acquisition of the fucoxanthin plastid was accompanied by loss of the ancestral peridinin plastid and all of its nuclear-encoded genes, likely because of incompatibilities in the biochemical environments between the two kinds of plastids.

Subsequent studies, however, showed that the newly acquired fucoxanthin plastid still uses proteins from the peridinin plastid. Phylogenetic analyses of the second plastid-targeted GAPDH isoform (encoded by *gapC1-pd*) clustered it with significant support within the clade of peridinin plastid-targeted proteins (Nosenko et al. 2006; Patron et al. 2006). The same was observed for glutamate-1-semialdehyde 2,1-aminomutase, but without clear bootstrap support. Extensive analyses of ESTs from another fucoxanthin dinoflagellate, *K. micrum*, identified two additional proteins (thylakoid-bound ascorbate peroxidase and phosphoribulokinase) with unambiguously supported peridinin plastid origins (Nosenko et al. 2006; Patron et al. 2006). Comparable evolutionary histories were suggested for three other proteins (ATP synthase gamma subunit, ferredoxin, and ribulose phosphate isomerase) but with weaker statistical support. Further evidence that the earlier presence of a peridinin plastid strongly impacted the haptophyte tertiary endosymbiosis was provided by Dorrell and Howe (2012) who demonstrated extensive editing of transcripts in the fucoxanthin plastid of *Karenia mikimotoi*. Because such editing is characteristic of peridinin plastids, but absent from haptophytes, this strongly suggests that fucoxanthin plastids use editing machinery left behind by the original peridinin plastid.

These cumulative data clearly show that the newly acquired fucoxanthin plastid uses some genes from its peridinin predecessor, while many others were lost or replaced by those from the haptophyte during the establishment of the tertiary endosymbiosis. These latter genes were acquired by endosymbiotic gene transfer and significantly remodeled the dinoflagellate proteome. In total almost 20 plastid related, haptophyte genes have been found in *Karenia* and *Karlodinium* (Nosenko

Fig. 6 (continued) have coexisted for some time, performing complementary roles for their dinoflagellate host. To avoid protein mistargeting, however, transit peptides in proteins targeted to the fucoxanthin plastid (F) diverged from those typical for peridinin plastid proteins through decreased alanine content, as well as loss of a positive net charge and FVAP-motif. E: Once the fucoxanthin plastid was fully established, it likely took over functions previously associated with the peridinin plastid resulting in the latter's complete disappearance

et al. 2006; Patron et al. 2006). They are involved in many different metabolic processes, including electron transport, carbohydrate metabolism, fatty acid synthesis, translation, and RNA binding.

An interesting scenario of plastid replacement in fucoxanthin dinoflagellates was formulated by Patron et al. (2006). Because only peridinin-plastid genes involved in nonphotosynthetic processes were found in fucoxanthin dinoflagellates, they postulated that the haptophyte photosynthetic endosymbiont was engulfed by a dinoflagellate host in which the peridinin plastid already had lost its photosynthetic activity (Fig. 6). The plastid still could have been involved in vital nonphotosynthetic functions, such as fatty acid and isoprenoid biosynthesis (McFadden 2011). The peculiar nature of targeting domains in proteins imported into fucoxanthin plastids suggest that the peridinin and fucoxanthin plastids coexisted for some time (Patron et al. 2006). They resemble neither transit peptides of their haptophyte ancestors nor those of the earlier peridinin plastids. In *K. micrum*, these peptides lack both the positive net charge and the high frequency of alanine residues typical for plastid transit peptides (Patron et al. 2006; Patron and Waller 2007). The FVAP-motif common in peridinin dinoflagellates, but not in haptophytes, was found only in one of the 26 *Karlodinium* proteins investigated (Patron et al. 2006; Patron and Waller 2007). Moreover, only two *Karlodinium* proteins contain a second hydrophobic region downstream of the transit peptide, a characteristic of most proteins transported into the peridinin plastid (Patron et al. 2006; Patron and Waller 2007). All these changes in the N-terminal targeting signals probably evolved to prevent mistargeting of haptophyte plastid proteins to the peridinin plastid (Patron et al. 2006). Later, once all essential plastid-directed proteins and their functions were transferred to the fucoxanthin plastid proteome, the peridinin plastid could be lost entirely.

Perspectives: Broader Significance of Tertiary Endosymbiosis

Tertiary endosymbioses are not unique to dinoflagellates. A similar range of such "endosymbioses in progress" has been described in other protozoans from a number of different phylogenetic lineages (for reviews, see Fehling et al. 2007; Stoecker et al. 2009; Johnson 2011b). As in dinoflagellates, there is great taxonomic diversity in the endosymbiont/plastid donors involved. Various rhizarian protists (e.g., radiolarians, foraminifers, and euglyphid amoebae) have developed symbioses not only with green and red algae but also with dinoflagellates, diatoms, and haptophytes (for details, see Fehling et al. 2007; Stoecker et al. 2009; Johnson 2011a). Some of them, such as *Bulimina*, *Elphidium*, and *Nonion* foraminiferans, extract and retain only the plastids from their diatom prey (Bernhard et al. 2001; Bernhard 2003; Correira and Lee 2002b). Other foraminiferans sometimes keep more than one symbiont; for example, *Amphistegina* maintains two (rarely three)

pennate diatom species and sometimes the green alga *Chlorella* as well (Lee et al. 1986; Lee and Correia 2005). The ciliates *Laboea*, *Strombidium*, and *Tontonia* have been found to take up tertiary plastids from a variety of algal prey, including euglenoids, haptophytes, cryptophytes, stramenopiles, and dinoflagellates (Laval-Peuto and Febvre 1986; Stoecker et al. 1988; Stoecker and Silver 1990). Ciliates also vary substantially in their reliance on kleptoplasty; however, plastid retention frequently is obligatory (Stoecker et al. 2009; Esteban et al. 2010). Some ciliates take up intact endosymbionts; for example, *Maristentor dinoferus* and *Euplotes uncinatus* maintain whole *Symbiodinium* dinoflagellates within their cells (Lobban et al. 2002, 2005). Others, like *Perispira ovum*, hold onto plastids and a partial collection of their prey's cellular components, in this case plastids, mitochondria, and paramylon from *Euglena proxima* (Johnson et al. 1995).

Various multicellular animals also have evolved endosymbioses with algae containing secondary plastids (Rumpho et al. 2011; see also Wägele and Martin 2013). Perhaps, the best known are reef-building (scleractinian) corals, which harbor symbiotic dinoflagellates from the genus *Symbiodinium* (known as zooxanthellae) that deliver energy for carbonate deposition required for coral reef formation (Muscatine 1990; Baker 2003). In fact, *Symbiodinium* appears to be particularly adept at forming such relationships; these symbionts have been identified in diverse animals including sponges (*Spongilla*), sea anemones (*Anthopleura*), bivalves (e.g., *Tridacna*), and gastropods (nudibranchs) (Venn et al. 2008). Some turbellarians are known to host *Amphidinium* dinoflagellates and diatoms, among other algae (Fehling et al. 2007). The most interesting and well-characterized examples of animal kleptoplasty is, however, one from the sea slug *Elysia chlorotica*, which extracts and keeps active plastids from its major stramenopile algal food source, *V. litorea*, for as long as 10 months (Rumpho et al. 2011; Wägele and Martin 2013).

Although we have highlighted only a fraction of known cases in this chapter, mainly from dinoflagellates, it is clear that tertiary endosymbioses are important and widespread phenomena in nature. They have contributed substantially to the evolution of key algal taxa and continue to be important as both ecological and evolutionary drivers of modern photosynthetic biodiversity. It is interesting to reflect on how an endosymbiotic process, one that likely began with a single primary endosymbiosis between a cyanobacterium and protist about 1,500 Mya, has been woven so intimately and inextricably through the fabric of eukaryotic life.

Acknowledgments We are grateful to Dr. P.J. Hansen for helpful comments. We also would like to thank Prof. W. Löffelhardt for inviting us to write this chapter and for his patience during its preparation. This work was supported by grant UMO-2011/01/N/NZ8/00150 to P. Gagat and P. Mackiewicz and Wrocław University grant 1069/S/KBEE/2012 to A. Bodył.

References

Adolf JE, Stoecker DK, Harding LW (2006) The balance of autotrophy and heterotrophy during mixotrophic growth of *Karlodinium micrum* (Dinophyceae). J Plankton Res 28:737–751

Agrawal S, Striepen B (2010) More membranes, more proteins: complex protein import mechanisms into secondary plastids. Protist 161:672–687

Archibald JM (2009) The puzzle of plastid evolution. Curr Biol 19:R81–R88

Bachvaroff TR, Concepcion GT, Rogers CR, Herman EM, Delwiche CF (2004) Dinoflagellate expressed sequence tag data indicate massive transfer of chloroplast genes to the nuclear genome. Protist 155:65–78

Bachvaroff TR, Puerta MVS, Delwiche CF (2005) Chlorophyll *c*-containing plastid relationships based on analyses of a multigene data set with all four chromalveolate lineages. Mol Biol Evol 22:1772–1782

Bachvaroff TR, Sanchez-Puerta MV, Delwiche CF (2006) Rate variation as a function of gene origin in plastid-derived genes of peridinin-containing dinoflagellates. J Mol Evol 62:42–52

Bachvaroff TR, Adolf JE, Squier AH, Harvey HR, Place AR (2008) Characterization and quantification of karlotoxins by liquid chromatography-mass spectrometry. Harmful Algae 7:473–484

Bachvaroff TR, Handy SM, Place AR, Delwiche CF (2011) Alveolate phylogeny inferred using concatenated ribosomal proteins. J Eukaryot Microbiol 58:223–233

Baker AC (2003) Flexibility and specifity in coral-algal symbiosis: diversity, ecology, and biogeography of *Symbiodinium*. Annu Rev Ecol Evol Syst 34:661–689

Bergholtz T, Daugbjerg N, Moestrup Ø, Fernández-Tejedor M (2006) On the identity of *Karlodinium veneficum* and description of *Karlodinium armiger* sp. nov. (Dinophyceae), based on light and electron microscopy, nuclear-encoded LSU rDNA, and pigment composition. J Phycol 42:170–193

Bernhard JM (2003) Potential symbionts in bathyal foraminifera. Science 299:861

Bernhard JM, Buck KR, Barry JP (2001) Monterey Bay cold-seep biota: assemblages, abundance, and ultrastructure of living foraminifera. Deep-Sea Res I Oceanogr Res Pap 48:2233–2249

Bjørnland T, Haxo FT, Liaaen-Jensen S (2003) Carotenoids of the Florida red tide dinoflagellate: *Karenia brevis*. Biochem Syst Ecol 31:1147–1162

Bodył A (1999) How have apicomplexan plastids evolved? A hypothesis. In: From symbiosis to eukaryotism. Proceedings of the international congress on endocytobiology, symbiosis and biomedicine, endocytobiology VII, 5–9 Apr 1998, Freiburg im Breisgau. University of Freiburg, University of Geneva, pp 327–340

Bodył A, Moszczyński K (2006) Did the peridinin plastid evolve through tertiary endosymbiosis? A hypothesis. Eur J Phycol 41:435–448

Bodył A, Stiller JW, Mackiewicz P (2009a) Chromalveolate plastids: direct descent or multiple endosymbioses? Trends Ecol Evol 24:119–121

Bodył A, Mackiewicz P, Stiller JW (2009b) Early steps in plastid evolution: current ideas and controversies. Bioessays 31:1219–1232

Brand LE, Campbell L, Bresnan E (2012) *Karenia*: the biology and ecology of a toxic genus. Harmful Algae 14:156–178

Bullmann L, Haarmann R, Mirus O, Bredemeier R, Hempel F, Maier UG, Schleiff E (2010) Filling the gap, evolutionarily conserved Omp85 in plastids of chromalveolates. J Biol Chem 285: 6848–6856

Cavalier-Smith T (1993a) Kingdom protozoa and its 18 phyla. Microbiol Rev 57:953–994

Cavalier-Smith T (1993b) The origin, losses and gains of chloroplasts. In: Lewin RA (ed) Origins of plastids: symbiogenesis, prochlorophytes, and the orgins of chloroplasts. Chapman and Hall, New York, pp 291–349

Cavalier-Smith T (1999) Principles of protein and lipid targeting in secondary symbiogenesis: euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree. J Eukaryot Microbiol 46:347–366

Cavalier-Smith T (2006) Origin of mitochondria by intracellular enslavement of a photosynthetic purple bacterium. Proc Biol Sci 273:1943–1952

Cavalier-Smith T, Lee JJ (1985) Protozoa as hosts for endosymbioses and the conversion of symbionts into organelles. J Eukaryot Microbiol 32:376–379

Chesnick JM, Cox ER (1989) Fertilization and zygote development in the binucleate dinoflagellate *Peridinium balticum* (Pyrrhophyta). Am J Bot 76:1060–1072

Chesnick JM, Morden CW, Schmieg AM (1996) Identity of the endosymbiont of *Peridinium foliaceum* (Pyrrhophyta): analysis of the *rbcLS* operon. J Phycol 32:850–857

Chesnick JM, Kooistra WHCF, Wellbrock U, Medlin LK (1997) Ribosomal RNA analysis indicates a benthic pennate diatom ancestry for the endosymbionts of the dinoflagellates *Peridinium foliaceum* and *Peridinium balticum* (Pyrrhophyta). J Eukaryot Microbiol 44:314–320

Correira MJ, Lee JJ (2002a) How long do the plastids retained by *Elphidium excavatum* (Terquem) last in their host? Symbiosis 32:27–37

Correira MJ, Lee JJ (2002b) Fine structure of the plastids retained by the foraminifer *Elphidium excavatum* (Terquem). Symbiosis 32:15–26

Curtis BA, Tanifuji G, Burki F, Gruber A, Irimia M, Maruyama S, Arias MC, Ball SG, Gile GH, Hirakawa Y, Hopkins JF, Kuo A, Rensing SA, Schmutz J, Symeonidi A, Elias M, Eveleigh RJ, Herman EK, Klute MJ, Nakayama T, Obornik M, Reyes-Prieto A, Armbrust EV, Aves SJ, Beiko RG, Coutinho P, Dacks JB, Durnford DG, Fast NM, Green BR, Grisdale CJ, Hempel F, Henrissat B, Hoppner MP, Ishida K, Kim E, Koreny L, Kroth PG, Liu Y, Malik SB, Maier UG, McRose D, Mock T, Neilson JA, Onodera NT, Poole AM, Pritham EJ, Richards TA, Rocap G, Roy SW, Sarai C, Schaack S, Shirato S, Slamovits CH, Spencer DF, Suzuki S, Worden AZ, Zauner S, Barry K, Bell C, Bharti AK, Crow JA, Grimwood J, Kramer R, Lindquist E, Lucas S, Salamov A, McFadden GI, Lane CE, Keeling PJ, Gray MW, Grigoriev IV, Archibald JM (2012) Algal genomes reveal evolutionary mosaicism and the fate of nucleomorphs. Nature 492:59–65

Dang Y, Green BR (2009) Long transcripts from dinoflagellate chloroplast minicircles suggest 'rolling circle' transcription. J Biol Chem 285:5196–5203

Danne JC, Gornik SG, Waller RF (2012) An assessment of vertical inheritance versus endosymbiont transfer of nucleus-encoded genes for mitochondrial proteins following tertiary endosymbiosis in *Karlodinium micrum*. Protist 163:76–90

Daugbjerg N, Hansen G, Larsen J, Moestrup Ø (2000) Phylogeny of some of the major genera of dinoflagellates based on ultrastructure and partial LSU rDNA sequence data, including the erection of three new genera of unarmoured dinoflagellates. Phycologia 39:302–317

de Salas MF, Bolch CJS, Botes L, Nash G, Wright SW, Hallegraeff GM (2003) *Takayama* gen. nov. (Gymnodiniales, Dinophyceae), a new genus of unarmored dinoflagellates with sigmoid apical grooves, including the description of two new species. J Phycol 39:1233–1246

Delwiche CF (2007) The origin and evolution of dinoflagellates. In: Falkowski PGK, Knoll AH (eds) Evolution of primary producers in the sea. Elsevier Academic, Beijing, pp 191–205

Dodge JD (1983) Dinoflagellates: investigation and phylogenetic speculation. Br Phycol J 18: 335–356

Dodge JD (1984) The functional and phylogenetic significance of dinoflagellate eyespots. Biosystems 16:259–267

Dodge JD, Crawford RM (1969) Observations on the fine structure of the eyespot and associated organelles in the dinoflagellate *Glenodinium foliaceum*. J Cell Sci 5:479–493

Dorrell RG, Howe CJ (2012) Functional remodeling of RNA processing in replacement chloroplasts by pathways retained from their predecessors. Proc Natl Acad Sci USA 109: 18879–18884

Durnford DG, Gray MW (2006) Analysis of *Euglena gracilis* plastid-targeted proteins reveals different classes of transit sequences. Eukaryot Cell 5:2079–2091

Eriksen NT, Hayes KC, Lewitus AJ (2002) Growth responses of the mixotrophic dinoflagellates, *Cryptoperidiniopsis* sp. and *Pfiesteria piscicida*, to light under prey-saturated conditions. Harmful Algae 1:191–203

Escalera L, Reguera B, Takishita K, Yoshimatsu S, Koike K, Koike K (2011) Cyanobacterial endosymbionts in the benthic dinoflagellate *Sinophysis canaliculata* (Dinophysiales, Dinophyceae). Protist 162:304–314

Eschbach S, Speth V, Hansmann P, Sitte P (1990) Freeze-fracture study of the single membrane between host cell and endocytobiont in the dinoflagellates *Glenodinium foliaceum* and *Peridinium balticum*. J Phycol 26:324–328

Espelund M, Minge MA, Gabrielsen TM, Nederbragt AJ, Shalchian-Tabrizi K, Otis C, Turmel M, Lemieux C, Jakobsen KS (2012) Genome fragmentation is not confined to the peridinin plastid in dinoflagellates. PLoS One 7:e38809

Esteban GF, Fenchel T, Finlay BJ (2010) Mixotrophy in ciliates. Protist 161:621–641

Farmer MA, Roberts KR (1990) Organelle loss in the endosymbiont of *Gymnodinium acidotum* (Dinophyceae). Protoplasma 153:178–185

Farnelid H, Tarangkoon W, Hansen G, Hansen P, Riemann L (2010) Putative N_2-fixing heterotrophic bacteria associated with dinoflagellate-cyanobacteria consortia in the low-nitrogen Indian Ocean. Aquat Microb Ecol 61:105–117

Fehling J, Stoecker D, Baldauf S (2007) Photosynthesis and the eukaryote tree of life. In: Falkowski PG, Knoll AH (eds) Evolution of primary producers in the sea. Elsevier, New York, pp 76–107

Fensome R, Taylor F, Norris G, Sarjeant W, Warton D, Williams G (1993) A classification of living and fossil dinoflagellates. In: Micropaleontology, Special Publication no. 7. American Museum of Natural History, New York, pp 1–351

Fields SD, Rhodes RG (1991) Ingestion and retention of *Chroomonas* spp. (Cryptophyceae) by *Gymnodinium acidotum* (Dinophyceae). J Phycol 27:525–529

Figueroa RI, Bravo I, Fraga S, Garcés E, Llaveria G (2009) The life history and cell cycle of *Kryptoperidinium foliaceum*, a dinoflagellate with two eukaryotic nuclei. Protist 160:285–300

Foster RA, Carpenter EJ, Bergman B (2006a) Unicellular cyanobionts in open ocean dinoflagellates, radiolarians, and tintinnids: ultrastructural characterization and immunolocalization of phycoerythrin and nitrogenase. J Phycol 42:453–463

Foster RA, Collier JL, Carpenter EJ (2006b) Reverse transcription PCR amplification of cyanobacterial symbiont 16S rRNA sequences from single non-photosynthetic eukaryotic marine planktonic host cells. J Phycol 42:243–250

Gabrielsen TM, Minge MA, Espelund M, Tooming-Klunderud A, Patil V, Nederbragt AJ, Otis C, Turmel M, Shalchian-Tabrizi K, Lemieux C, Jakobsen KS (2011) Genome evolution of a tertiary dinoflagellate plastid. PLoS One 6:e19132

Gaines G, Taylor FJR (1985) Form and function of the dinoflagellate transverse flagellum. J Protozool 32:290–296

Garate-Lizarraga I, Muneton-Gomez MD (2008) Bloom of *Peridinium quinquecorne* Abe, in La Ensenada de La Paz, Gulf of California (July 2003). Acta Bot Mex 83:33–47

Garcia-Cuetos L, Moestrup Ø, Hansen PJ, Daugbjerg N (2010) The toxic dinoflagellate *Dinophysis acuminata* harbors permanent chloroplasts of cryptomonad origin, not kleptochloroplasts. Harmful Algae 9:25–38

Gast RJ, Moran DM, Dennett MR, Caron DA (2007) Kleptoplasty in an Antarctic dinoflagellate: caught in evolutionary transition? Environ Microbiol 9:39–45

Gilson PR, Su V, Slamovits CH, Reith ME, Keeling PJ, McFadden GI (2006) Complete nucleotide sequence of the chlorarachniophyte nucleomorph: nature's smallest nucleus. Proc Natl Acad Sci USA 103:9566–9571

Gómez F (2012a) A quantitative review of the lifestyle, habitat and trophic diversity of dinoflagellates (Dinoflagellata, Alveolata). Syst Biodivers 10:267–275

Gómez F (2012b) A checklist and classification of living dinoflagellates (Dinoflagellata, Alveolata). CICIMAR Oceánides 27:65–140

Gómez F, López-García P, Nowaczyk A, Moreira D (2009) The crustacean parasites *Ellobiopsis* Caullery, 1910 and *Thalassomyces* Niezabitowski, 1913 form a monophyletic divergent clade within the Alveolata. Syst Parasitol 74:65–74

Gómez F, López-García P, Moreira D (2011) Molecular phylogeny of Dinophysoid dinoflagellates: the systematic position of *Oxyphysis oxytoxoides* and the *Dinophysis hastata* group (Dinophysales, Dinophyceae). J Phycol 47:393–406

Gould SB, Sommer MS, Kroth PG, Gile GH, Keeling PJ, Maier UG (2006) Nucleus-to-nucleus gene transfer and protein retargeting into a remnant cytoplasm of cryptophytes and diatoms. Mol Biol Evol 23:2413–2422

Green BR (2011) Chloroplast genomes of photosynthetic eukaryotes. Plant J 66:34–44

Green BJ, Fox TC, Rumpho ME (2005) Stability of isolated algal chloroplasts that participate in a unique mollusc/kleptoplast association. Symbiosis 40:31–40

Groisillier A, Massana R, Valentin KU, Vaulot D, Guillou L (2006) Genetic diversity and habitats of two enigmatic marine alveolate lineages. Aquat Microb Ecol 42:277–291

Grosche C, Hempel F, Bolte K, Abram L, Maier UG, Zauner S (2013) Protein import into complex plastids: current findings and perspectives. In: Löffelhardt W (ed) Endosymbiosis. Springer, Wien New York, pp 215–232

Gruber A, Vugrinec S, Hempel F, Gould SB, Maier UG, Kroth PG (2007) Protein targeting into complex diatom plastids: functional characterisation of a specific targeting motif. Plant Mol Biol 64:519–530

Grzymski J, Schofield OM, Falkowski PG, Bernhard JM (2002) The function of plastids in the deep-sea benthic foraminifer, *Nonionella stella*. Limnol Oceanogr 47:1569–1580

Guillou L, Viprey M, Chambouvet A, Welsh RM, Kirkham AR, Massana R, Scanlan DJ, Worden AZ (2008) Widespread occurrence and genetic diversity of marine parasitoids belonging to Syndiniales (Alveolata). Environ Microbiol 10:3349–3365

Gustafson DE, Stoecker DK, Johnson MD, Van Heukelem WF, Sneider K (2000) Cryptophyte algae are robbed of their organelles by the marine ciliate *Mesodinium rubrum*. Nature 405: 1049–1052

Hackett JD, Maranda L, Yoon HS, Bhattacharya D (2003) Phylogenetic evidence for the cryptophyte origin of the plastid of *Dinophysis* (Dinophysiales, Dinophyceae). J Phycol 39: 440–448

Hackett JD, Anderson DM, Erdner DL, Bhattacharya D (2004a) Dinoflagellates: a remarkable evolutionary experiment. Am J Bot 91:1523–1534

Hackett JD, Yoon HS, Soares MB, Bonaldo MF, Casavant TL, Scheetz TE, Nosenko T, Bhattacharya D (2004b) Migration of the plastid genome to the nucleus in a peridinin dinoflagellate. Curr Biol 14:213–218

Hallegraeff GM, Lucas IAN (1988) The marine dinoflagellate genus *Dinophysis* (Dinophyceae): photosynthetic, neritic and non-photosynthetic, oceanic species. Phycologia 27:25–42

Hansen PJ (2011) The role of photosynthesis and food uptake for the growth of marine mixotrophic dinoflagellates. J Eukaryot Microbiol 58:203–214

Hansen PJ, Calado AJ (1999) Phagotrophic mechanisms and prey selection in free-living dinoflagellates. J Eukaryot Microbiol 46:382–389

Hansen G, Daugbjerg N, Henriksen P (2000a) Comparative study of *Gymnodinium mikimotoi* and *Gymnodinium aureolum*, comb. nov. (=*Gyrodinium aureolum*) based on morphology, pigment composition, and molecular data. J Phycol 36:394–410

Hansen G, Moestrup Ø, Roberts KR (2000b) Light and electron microscopical observations on the type species of *Gymnodinium*, *G. fuscum* (*Dinophyceae*). Phycologia 39:365–376

Hansen PJ, Miranda L, Azanza R (2004) Green *Noctiluca scintillans*: a dinoflagellate with its own greenhouse. Mar Ecol Prog Ser 275:79–87

Häuber M, Müller S, Speth V, Maier U (1994) How to evolve a complex plastid? – A hypothesis. Bot Acta 107:383–386

Hempel F, Felsner G, Maier UG (2010) New mechanistic insights into pre-protein transport across the second outermost plastid membrane of diatoms. Mol Microbiol 76:793–801

Hewes CD, Mitchell BG, Moisan TA, Vernet M, Reid FMH (1998) The phycobilin signatures of chloroplasts from three dinoflagellate species: a microanalytical study of *Dinophysis caudata*, *D. fortii*, and *D. acuminata* (Dinophysiales, Dinophyceae). J Phycol 34:945–951

Hirakawa Y, Gile GH, Ota S, Keeling PJ, Ishida K-I (2010) Characterization of periplastidal compartment-targeting signals in chlorarachniophytes. Mol Biol Evol 27:1538–1545

Hirakawa Y, Burki F, Keeling PJ (2011) Nucleus- and nucleomorph-targeted histone proteins in a chlorarachniophyte alga. Mol Microbiol 80:1439–1449

Hoppenrath M, Leander BS (2010) Dinoflagellate phylogeny as inferred from heat shock protein 90 and ribosomal gene sequences. PLoS One 5:e13220

Horiguchi T (2004) Origin and evolution of dinoflagellates with a diatom endosymbiont. In: Proceedings of international symposium on "dawn of a new natural history – integration of geoscience and biodiversity studies", neo-science of natural history: integration of geoscience and biodiversity studies, 5–6 Mar 2004, Sapporo. Hokkaido University, Hokkaido, pp 53–59

Horiguchi T (2006) Algae and their chloroplasts with particular reference to the dinoflagellates. Paleontol Res 10:299–309

Horiguchi T, Pienaar RN (1992) *Amphidinium latum* Lebour (Dinophyceae), a sand-dwelling dinoflagellate feeding on cryptomonads. Jpn J Phycol 40:353–363

Horiguchi T, Pienaar RN (1994a) Ultrastructure of a new marine sand-dwelling dinoflagellate, *Gymnodinium quadrilobatum* sp. nov. (Dinophyceae) with special reference to its endosymbiotic alga. Eur J Phycol 29:237–245

Horiguchi T, Pienaar RN (1994b) Ultrastructure and ontogeny of a new type of eyespot in dinoflagellates. Protoplasma 179:142–150

Horiguchi T, Takano Y (2006) Serial replacement of a diatom endosymbiont in the marine dinoflagellate *Peridinium quinquecorne* (Peridiniales, Dinophyceae). Phycol Res 54:193–200

Howe CJ, Nisbet RE, Barbrook AC (2008) The remarkable chloroplast genome of dinoflagellates. J Exp Bot 59:1035–1045

Imanian B, Keeling PJ (2007) The dinoflagellates *Durinskia baltica* and *Kryptoperidinium foliaceum* retain functionally overlapping mitochondria from two evolutionarily distinct lineages. BMC Evol Biol 7:172

Imanian B, Pombert JF, Keeling PJ (2010) The complete plastid genomes of the two 'dinotoms' *Durinskia baltica* and *Kryptoperidinium foliaceum*. PLoS One 5:e10711

Imanian B, Pombert JF, Dorrell RG, Burki F, Keeling PJ (2012) Tertiary endosymbiosis in two dinotoms has generated little change in the mitochondrial genomes of their dinoflagellate hosts and diatom endosymbionts. PLoS One 7:e43763

Inaba T, Schnell DJ (2008) Protein trafficking to plastids: one theme, many variations. Biochem J 413:15–28

Inagaki Y, Roger AJ (2006) Phylogenetic estimation under codon models can be biased by codon usage heterogeneity. Mol Phylogenet Evol 40:428–434

Inagaki Y, Dacks JB, Doolittle WF, Watanabe KI, Ohama T (2000) Evolutionary relationship between dinoflagellates bearing obligate diatom endosymbionts: insight into tertiary endosymbiosis. Int J Syst Evol Microbiol 50:2075–2081

Inagaki Y, Simpson AG, Dacks JB, Roger AJ (2004) Phylogenetic artifacts can be caused by leucine, serine, and arginine codon usage heterogeneity: dinoflagellate plastid origins as a case study. Syst Biol 53:582–593

Ishida K, Green BR (2002) Second- and third-hand chloroplasts in dinoflagellates: phylogeny of oxygen-evolving enhancer 1 (PsbO) protein reveals replacement of a nuclear-encoded plastid gene by that of a haptophyte tertiary endosymbiont. Proc Natl Acad Sci USA 99:9294–9299

Jacobson DM, Andersen RA (1994) The discovery of mixotrophy in photosynthetic species of *Dinophysis* (Dinophyceae): light and electron microscopical observations of food vacuoles in *Dinophysis acuminata, D. norvegica* and two heterotrophic dinophysoid dinoflagellates. Phycologia 33:97–110

Jakobsen H, Hansen P, Larsen J (2000) Growth and grazing responses of two chloroplast-retaining dinoflagellates: effect of irradiance and prey species. Mar Ecol Prog Ser 201:121–128

Janouskovec J, Horak A, Obornik M, Lukes J, Keeling PJ (2010) A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids. Proc Natl Acad Sci USA 107: 10949–10954

Janson S (2004) Molecular evidence that plastids in the toxin-producing dinoflagellate genus *Dinophysis* originate from the free-living cryptophyte *Teleaulax amphioxeia*. Environ Microbiol 6:1102–1106

Janson S, Granéli E (2003) Genetic analysis of the *psbA* gene from single cells indicates a cryptomonad origin of the plastid in *Dinophysis* (Dinophyceae). Phycologia 42:473–477

Jeffrey SW, Vesk M (1976) Further evidence for a membrane-bound endosymbiont within the dinoflagellate *Peridinium foliaceum*. J Phycol 12:450–455

Jensen MH, Daugbjerg N (2009) Molecular phylogeny of selected species of the order Dinophysiales (Dinophyceae) – testing the hypothesis of a dinophysioid radiation. J Phycol 45:1136–1152

Jeong HJ, Du Yoo Y, Kim JS, Seong KA, Kang NS, Kim TH (2010) Growth, feeding and ecological roles of the mixotrophic and heterotrophic dinoflagellates in marine planktonic food webs. Ocean Sci J 45:65–91

Johnson MD (2011a) The acquisition of phototrophy: adaptive strategies of hosting endosymbionts and organelles. Photosynth Res 107:117–132

Johnson MD (2011b) Acquired phototrophy in ciliates: a review of cellular interactions and structural adaptations. J Eukaryot Microbiol 58:185–195

Johnson PW, Donaghay PL, Small EB, Sieburth JM (1995) Ultrastructure and ecology of *Perispira ovum* (Ciliophora: Litostomatea): an aerobic, planktonic ciliate that sequesters the chloroplasts, mitochondria, and paramylon of *Euglena proxima* in a micro-oxic habitat. J Eukaryot Microbiol 42:323–335

Johnson MD, Tengs T, Oldach D, Stoecker DK (2006) Sequestration, performance, and functional control of cryptophyte plastids in the ciliate *Myrionecta rubra* (Ciliophora). J Phycol 42: 1235–1246

Johnson MD, Oldach D, Delwiche CF, Stoecker DK (2007) Retention of transcriptionally active cryptophyte nuclei by the ciliate *Myrionecta rubra*. Nature 445:426–428

Keeling PJ (2009) Chromalveolates and the evolution of plastids by secondary endosymbiosis. J Eukaryot Microbiol 56:1–8

Keeling PJ (2010) The endosymbiotic origin, diversification and fate of plastids. Philos Trans R Soc Lond B Biol Sci 365:729–748

Kempton JW, Wolny J, Tengs T, Rizzo P, Morris R, Tunnell J, Scott P, Steidinger K, Hymel SN, Lewitus AJ (2002) *Kryptoperidinium foliaceum* blooms in South Carolina: a multi-analytical approach to identification. Harmful Algae 1:383–392

Kim E, Archibald J (2010) Plastid evolution: gene transfer and the maintenance of 'stolen' organelles. BMC Biol 8:73

Kim M, Nam SW, Shin W, Coats DW, Park MG (2012) *Dinophysis caudata* (Dinophyceae) sequesters and retains plastids from the mixotrophic ciliate prey *Mesodinium rubrum*. J Phycol 48:569–579

Kite GC, Dodge JD (1985) Structural organization of plastid DNA in two anomalously pigmented dinoflagellates. J Phycol 21:50–56

Kite G, Dodge J (1988) Cell and chloroplast ultrastructure in *Gyrodinium aureolum* and *Gymnodinium galatheanum*. Two marine dinoflagellates containing an unusual carotenoid. Sarsia 73:131–138

Kite GC, Rothschild LJ, Dodge JD (1988) Nuclear and plastid DNAs from the binucleate dinoflagellates *Glenodinium* (*Peridinium*) *foliaceum* and *Peridinium balticum*. Biosystems 21:151–163

Koike K, Takishita K (2008) Anucleated cryptophyte vestiges in the gonyaulacalean dinoflagellates *Amylax buxus* and *Amylax triacantha* (Dinophyceae). Phycol Res 56:301–311

Koike K, Koike K, Takagi M, Ogata T, Ishimaru T (2000) Evidence of phagotrophy in *Dinophysis fortii* (Dinophysiales, Dinophyceae), a dinoflagellate that causes diarrhetic shellfish poisoning (DSP). Phycol Res 48:121–124

Koike K, Sekiguchi H, Kobiyama A, Takishita K, Kawachi M, Koike K, Ogata T (2005) A novel type of kleptoplastidy in *Dinophysis* (Dinophyceae): presence of haptophyte-type plastid in *Dinophysis mitra*. Protist 156:225–237

Kreimer G (1999) Reflective properties of different eyespot types in dinoflagellates. Protist 150: 311–323

Kwok AC, Wong JT (2003) Cellulose synthesis is coupled to cell cycle progression at G1 in the dinoflagellate *Crypthecodinium cohnii*. Plant Physiol 131:1681–1691

Lane CE, van den Heuvel K, Kozera C, Curtis BA, Parsons BJ, Bowman S, Archibald JM (2007) Nucleomorph genome of *Hemiselmis andersenii* reveals complete intron loss and compaction as a driver of protein structure and function. Proc Natl Acad Sci USA 104:19908–19913

Larsen J (1988) An ultrastructural study of *Amphidinium poecilochroum* (Dinophyceae), a phagotrophic dinoflagellate feeding on small species of cryptophytes. Phycologia 27:366–377

Laval-Peuto M, Febvre M (1986) On plastid symbiosis in *Tontonia appendiculariformis* (Ciliophora, Oligotrichina). Biosystems 19:137–158

Leander BS, Keeling PJ (2003) Morphostasis in alveolate evolution. Trends Ecol Evol 18:395–402

Lee JJ, Correia M (2005) Endosymbiotic diatoms from previously unsampled habitats. Symbiosis 38:251–260

Lee JJ, Erez J, McEnery ME, Lagziel A, Xenophontos X (1986) Experiments on persistence of endosymbiotic diatoms in the larger foraminifer: *Amphistegina lessonii*. Symbiosis 1:211–226

Lewitus AJ, Glasgow HB, Burkholder JM (1999) Kleptoplastidy in the toxic dinoflagellate *Pfiesteria piscicida* (Dinophyceae). J Phycol 35:303–312

Lin S (2011) Genomic understanding of dinoflagellates. Res Microbiol 162:551–569

Lin S, Zhang H, Spencer D, Norman J, Gray M (2002) Widespread and extensive editing of mitochondrial mRNAs in dinoflagellates. J Mol Biol 320:727–739

Lin S, Zhang H, Gray MW (2008) RNA editing in dinoflagellates and its implications for the evolutionary history of the editing machinery. In: Smith H (ed) RNA and DNA editing: molecular mechanisms and their integration into biological systems. Wiley, Hoboken, NJ, pp 280–309

Linares M, Carter D, Gould SB (2013) *Chromera* et al: Novel photosynthetic alveolates and apicomplexan relatives. In: Löffelhardt W (ed) Endosymbiosis. Springer, Wien New York, pp 183–196

Lobban C, Schefter M, Simpson A, Pochon X, Pawlowski J, Foissner W (2002) *Maristentor dinoferus* n. gen., n. sp., a giant heterotrich ciliate (Spirotrichea: Heterotrichida) with zooxanthellae, from coral reefs on Guam, Mariana Islands. Mar Biol 141:207–208

Lobban C, Modeo L, Verni F, Rosati G (2005) *Euplotes uncinatus* (Ciliophora, Hypotrichia), a new species with zooxanthellae. Mar Biol 147:1055–1061

Lowe CD, Keeling PJ, Martin LE, Slamovits CH, Watts PC, Montagnes DJ (2011) Who is *Oxyrrhis marina*? Morphological and phylogenetic studies on an unusual dinoflagellate. J Plankton Res 33:555–567

Löffelhardt W (2013) The single primary endosymbiotic event leading to the Archaeplastida. In: Löffelhardt W (ed) Endosymbiosis. Springer, Wien New York, pp 39–52

Lucas IAN (1991) Symbionts of the tropical Dinophysiales (Dinophyceae). Ophelia 33:213–224

Lucas IAN, Vesk M (1990) The fine structure of two photosynthetic species of *Dinophysis* (Dinophysiales, Dinophyceae). J Phycol 26:345–357

Marasigan AN, Sato S, Fukuyo Y, Kodama M (2001) Accumulation of a high level of diarrhetic shellfish toxins in the green mussel *Perna viridis* during a bloom of *Dinophysis caudata* and *Dinophysis miles* in Sapian Bay, Panay Island, the Philippines. Fish Sci 67:994–996

McEwan ML, Keeling PJ (2004) HSP90, tubulin and actin are retained in the tertiary endosymbiont genome of *Kryptoperidinium foliaceum*. J Eukaryot Microbiol 51:651–659

McFadden GI (2011) The apicoplast. Protoplasma 248:641–650

Meyer-Harms B, Pollehne F (1998) Alloxanthin in *Dinophysis norvegica* (Dinophysiales, Dinophyceae) from the baltic sea. J Phycol 34:280–285

Minge MA, Shalchian-Tabrizi K, Torresen OK, Takishita K, Probert I, Inagaki Y, Klaveness D, Jakobsen KS (2010) A phylogenetic mosaic plastid proteome and unusual plastid-targeting signals in the green-colored dinoflagellate *Lepidodinium chlorophorum*. BMC Evol Biol 10: 191

Minnhagen S, Janson S (2006) Genetic analyses of *Dinophysis* spp. support kleptoplastidy. FEMS Microbiol Ecol 57:47–54

Minnhagen S, Carvalho WF, Salomon PS, Janson S (2008) Chloroplast DNA content in *Dinophysis* (Dinophyceae) from different cell cycle stages is consistent with kleptoplasty. Environ Microbiol 10:2411–2417

Moestrup Ø, Daugbjerg N (2007) On dinoflagellate phylogeny and classification. In: Brodie J, Lewis J (eds) Unraveling the algae: the past, present, and future of algal systematics. CRC, New York, pp 215–230

Moore RB, Obornik M, Janouskovec J, Chrudimsky T, Vancova M, Green DH, Wright SW, Davies NW, Bolch CJ, Heimann K, Slapeta J, Hoegh-Guldberg O, Logsdon JM, Carter DA (2008) A photosynthetic alveolate closely related to apicomplexan parasites. Nature 451: 959–963

Moreira D, López-Garcıa P (2002) The molecular ecology of microbial eukaryotes unveils a hidden world. Trends Microbiol 10:31–38

Morris RL, Fuller CB, Rizzo PJ (1993) Nuclear basic proteins from the binucleate dinoflagellate *Peridinium foliaceum* (Pyrrophyta). J Phycol 29:342–347

Muscatine L (1990) The role of symbiotic algae in carbon and energy flux in reef corals. In: Dubinsky Z (ed) Coral reefs. Elsevier, Amsterdam, pp 75–87

Nagai S, Nitshitani G, Tomaru Y, Sakiyama S (2008) Predation by the toxic dinoflagellate *Dinophysis fortii* on the ciliate *Myrionecta rubra* and observation of sequestration of ciliate chloroplast. J Phycol 44:909–922

Nassoury N, Cappadocia M, Morse D (2003) Plastid ultrastructure defines the protein import pathway in dinoflagellates. J Cell Sci 116:2867–2874

Nielsen L, Krock B, Hansen P (2012) Effects of light and food availability on toxin production, growth and photosynthesis in *Dinophysis acuminata*. Mar Ecol Prog Ser 471:37–50

Nishitani G, Miyamura K, Imai I (2003) Trying to cultivation of *Dinophysis caudata* (Dinophyceae) and the appearance of small cells. Plankton Biol Ecol 50:31–36

Nishitani G, Nagai S, Sakiyama S, Kamiyama T (2008a) Successful cultivation of the toxic dinoflagellate *Dinophysis caudata* (Dinophyceae). Plankton Benthos Res 3:78–85

Nishitani G, Nagai S, Takano Y, Sakiyama S, Baba K, Kamiyama T (2008b) Growth characteristics and phylogenetic analysis of the marine dinoflagellate *Dinophysis infundibulus* (Dinophyceae). Aquat Microb Ecol 52:209–221

Nishitani G, Nagai S, Baba K, Kiyokawa S, Kosaka Y, Miyamura K, Nishikawa T, Sakurada K, Shinada A, Kamiyama T (2010) High-level congruence of *Myrionecta rubra* prey and *Dinophysis* species plastid identities as revealed by genetic analyses of isolates from Japanese coastal waters. Appl Environ Microbiol 76:2791–2798

Nishitani G, Nagai S, Hayakawa S, Kosaka Y, Sakurada K, Kamiyama T, Gojobori T (2012) Multiple plastids collected by the dinoflagellate *Dinophysis mitra* through kleptoplastidy. Appl Environ Microbiol 78:813–821

Nosenko T, Lidie KL, Van Dolah FM, Lindquist E, Cheng JF, Bhattacharya D (2006) Chimeric plastid proteome in the Florida "red tide" dinoflagellate *Karenia brevis*. Mol Biol Evol 23: 2026–2038

Nowack ECM, Melkonian M (2010) Endosymbiotic associations within protists. Philos Trans R Soc Lond B Biol Sci 365:699–712

Oakley BR, Taylor FJR (1978) Evidence for a new type of endosymbiotic organization in a population of the ciliate *Mesodinium rubrum* from British Columbia. Biosystems 10:361–369

Orr RJS, Murray SA, Stüken A, Rhodes L, Jakobsen KS (2012) When naked became armored: an eight-gene phylogeny reveals monophyletic origin of theca in dinoflagellates. PLoS One 7: e50004

Oudot-Le Secq MP, Grimwood J, Shapiro H, Armbrust EV, Bowler C, Green BR (2007) Chloroplast genomes of the diatoms *Phaeodactylum tricornutum* and *Thalassiosira pseudonana*: comparison with other plastid genomes of the red lineage. Mol Genet Genomics 277:427–439

Park E, Rapoport TA (2012) Mechanisms of Sec61/SecY-mediated protein translocation across membranes. Annu Rev Biophys 41:21–40

Park MG, Kim S, Kim HS, Myung G, Kang YG, Yih W (2006) First successful culture of the marine dinoflagellate *Dinophysis acuminata*. Aquat Microb Ecol 45:w101–106

Park MG, Park JS, Kim M, Yih W (2008) Plastid dynamics during survival of *Dinophysis caudata* without its ciliate prey. J Phycol 44:1154–1163

Park MG, Kim M, Kim S, Yih W (2010) Does *Dinophysis caudata* (Dinophyceae) have permanent plastids? J Phycol 46:236–242

Patron NJ, Waller RF (2007) Transit peptide diversity and divergence: a global analysis of plastid targeting signals. Bioessays 29:1048–1058

Patron NJ, Waller RF, Archibald JM, Keeling PJ (2005) Complex protein targeting to dinoflagellate plastids. J Mol Biol 348:1015–1024

Patron NJ, Waller RF, Keeling PJ (2006) A tertiary plastid uses genes from two endosymbionts. J Mol Biol 357:1373–1382

Pienaar RN, Sakai H, Horiguchi T (2007) Description of a new dinoflagellate with a diatom endosymbiont, *Durinskia capensis* sp. nov. (Peridiniales, Dinophyceae) from South Africa. J Plant Res 120:247–258

Pozdnyakov I, Skarlato S (2012) Dinoflagellate amphiesma at different stages of the life cycle. Protistology 7:108–115

Qiu D, Huang L, Liu S, Lin S (2011) Nuclear, mitochondrial and plastid gene phylogenies of *Dinophysis miles* (Dinophyceae): evidence of variable types of chloroplasts. PLoS One 6: e29398

Reguera B, Velo-Suárez L, Raine R, Park MG (2012) Harmful *Dinophysis* species: a review. Harmful Algae 14:87–106

Reinbothe S, Quigley F, Springer A, Schemenewitz A, Reinbothe C (2004) The outer plastid envelope protein Oep16: role as precursor translocase in import of protochlorophyllide oxidoreductase A. Proc Natl Acad Sci USA 101:2203–2208

Richly E, Leister D (2004) An improved prediction of chloroplast proteins reveals diversities and commonalities in the chloroplast proteomes of *Arabidopsis* and rice. Gene 329:11–16

Rumpho ME, Pelletreau KN, Moustafa A, Bhattacharya D (2011) The making of a photosynthetic animal. J Exp Biol 214:303–311

Saldarriaga JF, McEwan ML, Fast NM, Taylor FJ, Keeling PJ (2003) Multiple protein phylogenies show that *Oxyrrhis marina* and *Perkinsus marinus* are early branches of the dinoflagellate lineage. Int J Syst Evol Microbiol 53:355–365

Sampayo MAM (1993) Trying to cultivate *Dinophysis* spp. In: Smayda TJ, Shimizu Y (eds) Toxic phytoplankton blooms in the sea. Elsevier Science, New York, pp 807–810

Sanchez Puerta MV, Bachvaroff TR, Delwiche CF (2005) The complete plastid genome sequence of the haptophyte *Emiliania huxleyi*: a comparison to other plastid genomes. DNA Res 12: 151–156

Sanchez-Puerta MV, Lippmeier JC, Apt KE, Delwiche CF (2007a) Plastid genes in a non-photosynthetic dinoflagellate. Protist 158:105–117

Sanchez-Puerta M, Bachvaroff TR, Delwiche CF (2007b) Sorting wheat from chaff in multi-gene analyses of chlorophyll *c*-containing plastids. Mol Biol Evol 44:885–897

Schnepf E (1993) From prey via endosymbiont to plastid: comparative studies in dinoflagellates. In: Lewin R (ed) Origins of plastids. Chapman & Hall, London, pp 53–76

Schnepf E (2004) Protoctists and microalgae: antagonistic and mutualistic associations and the symbiosis of plastids. In: Esser K, Lüttge U, Beyschlag W, Murata J (eds) Progress in Botany, vol 65. Springer, Heidelberg, pp 3–51

Schnepf E, Elbrächter M (1988) Cryptophycean-like double-membrane-bound chloroplast in the dinoflagellate, *Dynophysis* Ehrenb.: evolutionary, phylogenetic and toxicological implications. Bot Acta 101:196–203

Schnepf E, Elbrächter M (1999) Dinophyte chloroplasts and phylogeny – a review. Grana 38: 81–97

Schnepf E, Winter S, Mollenhauer D (1989) *Gymnodinium aeruginosum* (Dinophyta): a blue-green dinoflagellate with a vestigial, anucleate, cryptophycean endosymbiont. Plant Syst Evol 164:75–91

Schweikert M, Elbrächter M (2004) First ultrastructural investigations of the consortium between a phototrophic eukaryotic endocytobiont and *Podolampas bipes* (Dinophyceae). Phycologia 43:614–623

Shalchian-Tabrizi K, Skanseng M, Ronquist F, Klaveness D, Bachvaroff TR, Delwiche CF, Botnen A, Tengs T, Jakobsen KS (2006) Heterotachy processes in rhodophyte-derived secondhand plastid genes: implications for addressing the origin and evolution of dinoflagellate plastids. Mol Biol Evol 23:1504–1515

Skovgaard A (1998) Role of chloroplast retention in a marine dinoflagellate. Aquat Microb Ecol 15:293–301

Skovgaard A (2000) A phagotrophically derivable growth factor in the plastidic dinoflagellate *Gyrodinium resplendens* (Dinophyceae). J Phycol 36:1069–1078

Slamovits CH, Keeling PJ (2008) Plastid-derived genes in the nonphotosynthetic alveolate *Oxyrrhis marina*. Mol Biol Evol 25:1297–1306

Sláviková S, Vacula R, Fang Z, Ehara T, Osafune T, Schwartzbach SD (2005) Homologous and heterologous reconstitution of Golgi to chloroplast transport and protein import into the complex chloroplasts of *Euglena*. J Cell Sci 118:1651–1661

Sommer MS, Schleiff E (2013) Evolution of the protein translocon at the envelopes of chloroplasts. In: Löffelhardt W (ed) Endosymbiosis. Springer, Wien New York, pp 81–110

Steidinger KA, Truby EW, Dawes CJ (1978) Ultrastructure of the red tide dinoflagellate *Gymnodinium breve*. I. General description. J Phycol 14:72–79

Steidinger KA, Landsberg JH, Truby EW, Roberts BS (1998) First report of *Gymnodinium pulchellum* (Dinophyceae) in North America and associated fish kills in the Indian River, Florida. J Phycol 34:431–437

Stiller W, Reel D, Johnson J (2002) The case for a single plastid origin revisited: convergent evolution in organellar gene content. J Phycol 38:34–34

Stoebe B, Maier U-G (2002) One, two, three: nature's tool box for building plastids. Protoplasma 219:123–130

Stoecker DK (1999) Mixotrophy among dinoflagellates. J Eukaryot Microbiol 46:397–401

Stoecker DK, Silver MW (1990) Replacement and aging of chloroplasts in *Strombidium capitatum* (Ciliophora: Oligotrichida). Mar Biol 107:491–502

Stoecker DK, Silver MW, Michaels AE, Davis LH (1988) Obligate mixotrophy in *Laboea strobila*, a ciliate which retains chloroplasts. Mar Biol 99:415–423

Stoecker DK, Johnson MD, de Vargas C, Not F (2009) Acquired phototrophy in aquatic protists. Aquat Microb Ecol 57:279–310

Sulli C, Fang Z, Muchhal U, Schwartzbach SD (1999) Topology of *Euglena* chloroplast protein precursors within endoplasmic reticulum to Golgi to chloroplast transport vesicles. J Biol Chem 274:457–463

Sweeney BM (1978) Ultrastructure of *Noctiluca miliaris* (Pyrrophyta) with green flagellate symbionts. J Phycol 14:116–120

Takano Y, Hansen G, Fujita D, Horiguchi T (2008) Serial replacement of diatom endosymbionts in two freshwater dinoflagellates, *Peridiniopsis* spp. (Peridiniales, Dinophyceae). Phycologia 47:41–53

Takishita K, Koike K, Maruyama T, Ogata T (2002) Molecular evidence for plastid robbery (kleptoplastidy) in *Dinophysis*, a dinoflagellate causing diarrhetic shellfish poisoning. Protist 153:293–302

Takishita K, Ishida K-I, Maruyama T (2004) Phylogeny of nuclear-encoded plastid-targeted GAPDH gene supports separate origins for the peridinin- and the fucoxanthin derivative-containing plastids of dinoflagellates. Protist 155:447–458
Tamura M, Shimada S, Horiguchi T (2005) *Galeidiniium rugatum* gen. et sp. nov. (Dinophyceae), a new coccoid dinoflagellate with a diatom endosymbiont. J Phycol 41:658–671
Tanaka T, Fukuda Y, Yoshino T, Maeda Y, Muto M, Matsumoto M, Mayama S, Matsunaga T (2011) High-throughput pyrosequencing of the chloroplast genome of a highly neutral-lipid-producing marine pennate diatom, *Fistulifera* sp. strain JPCC DA0580. Photosynth Res 109: 223–229
Tanifuji G, Archibald JM (2013) Nucleomorph comparative genomics. In: Löffelhardt W (ed) Endosymbiosis. Springer, Wien New York, pp 197–214
Tarangkoon W, Hansen G, Hansen PJ (2007) Dinoflagellate/cyanobacteria consortia in the tropical Indian Ocean and the north west Australian Sea. J Phycol 43:38
Taylor FJR (1976) Dinoflagellates from the International Indian Ocean expedition. A report on material collected by the R.V. "Anton Bruun" 1963–1964. Bibliotheca Bot 132:1–234
Taylor FJRM (2004) Illumination or confusion? Dinoflagellate molecular phylogenetic data viewed from a primarily morphological standpoint. Phycol Res 52:308–324
Taylor FJR, Hoppenrath M, Saldarriaga JF (2007) Dinoflagellate diversity and distribution. Biodivers Conserv 17:407–418
Tengs T, Dahlberg OJ, Shalchian-Tabrizi K, Klaveness D, Rudi K, Delwiche CF, Jakobsen KS (2000) Phylogenetic analyses indicate that the 19' hexanoyloxy-fucoxanthin-containing dinoflagellates have tertiary plastids of haptophyte origin. Mol Biol Evol 17:718–729
Tippit DH, Pickett-Heaps JD (1976) Apparent amitosis in the binucleate dinoflagellate *Peridinium balticum*. J Cell Sci 21:273–289
Tomas RN, Cox ER (1973) Observations on the symbiosis of *Peridinium balticum* and its intracellular alga. I. Ultrastructure. J Phycol 9:304–323
van Wijk KJ (2004) Plastid proteomics. Plant Physiol Biochem 42:963–977
Venn AA, Loram JE, Douglas AE (2008) Photosynthetic symbioses in animals. J Exp Bot 59: 1069–1080
Vesk M, Dibbayawan TP, Vesk PA (1996) Immunogold localization of phycoerythrin in chloroplasts of *Dinophysis acuminata* and *D. fortii* (Dinophysiales, Dinophyta). Phycologia 35:234–238
Wägele H, Martin WF (2013) Endosymbioses in sacoglossan seaslugs: plastid-bearing animals that keep photosynthetic organelles without borrowing genes. In: Löffelhardt W (ed) Endosymbiosis. Springer, Wien New York, pp 291–324
Waller RF, Jackson CJ (2009) Dinoflagellate mitochondrial genomes: stretching the rules of molecular biology. Bioessays 31:237–245
Wang DZ (2008) Neurotoxins from marine dinoflagellates: a brief review. Mar Drugs 6:349–371
Wang Y, Morse D (2006) The plastid-encoded *psbA* gene in the dinoflagellate *Gonyaulax* is not encoded on a minicircle. Gene 371:206–210
Wang Y, Joly S, Morse D (2008) Phylogeny of dinoflagellate plastid genes recently transferred to the nucleus supports a common ancestry with red algal plastid genes. J Mol Evol 66:175–184
Watkins SM, Reich A, Fleming LE, Hammond R (2008) Neurotoxic shellfish poisoning. Mar Drugs 6:431–455
Whitney SM, Shaw DC, Yellowlees D (1995) Evidence that some dinoflagellates contain a ribulose-1, 5-bisphosphate carboxylase/oxygenase related to that of the alpha-proteobacteria. Proc Biol Sci 259:271–275
Wilcox LW, Wedemayer GJ (1984) *Gymnodinium acidotum* Nygaard (Pyrrophyta), a dinoflagellate with an endosymbiotic cryptomonad. J Phycol 20:236–242
Wilcox LW, Wedemayer GJ (1985) Dinoflagellate with blue-green chloroplasts derived from an endosymbiotic eukaryote. Science 227:192–194

Wisecaver J, Hackett J (2010) Transcriptome analysis reveals nuclear-encoded proteins for the maintenance of temporary plastids in the dinoflagellate *Dinophysis acuminata*. BMC Genomics 11:366

Wisecaver JH, Hackett JD (2011) Dinoflagellate genome evolution. Annu Rev Microbiol 65: 369–387

Withers NW, Cox ER, Tomas R, Haxo FT (1977) Pigments of the dinoflagellate *Peridinium balticum* and its photosynthetic endosymbiont. J Phycol 13:354–358

Yamaguchi H, Nakayama T, Kai A, Inouye I (2011) Taxonomy and phylogeny of a new kleptoplastidal dinoflagellate, *Gymnodinium myriopyrenoides* sp. nov. (Gymnodiniales, Dinophyceae), and its cryptophyte symbiont. Protist 162:650–667

Yoon HS, Hackett JD, Bhattacharya D (2002) A single origin of the peridinin- and fucoxanthin-containing plastids in dinoflagellates through tertiary endosymbiosis. Proc Natl Acad Sci USA 99:11724–11729

Yoon HS, Hackett JD, Van Dolah FM, Nosenko T, Lidie KL, Bhattacharya D (2005) Tertiary endosymbiosis driven genome evolution in dinoflagellate algae. Mol Biol Evol 22:1299–1308

Zapata M, Jeffrey S, Wright SW, Rodríguez F, Garrido JL, Clementson L (2004) Photosynthetic pigments in 37 species (65 strains) of Haptophyta: implications for oceanography and chemotaxonomy. Mar Ecol Prog Ser 270:83–102

Zauner S, Greilinger D, Laatsch T, Kowallik KV, Maier UG (2004) Substitutional editing of transcripts from genes of cyanobacterial origin in the dinoflagellate *Ceratium horridum*. FEBS Lett 577:535–538

Zhang Z, Green BR, Cavalier-Smith T (2000) Phylogeny of ultra-rapidly evolving dinoflagellate chloroplast genes: a possible common origin for sporozoan and dinoflagellate plastids. J Mol Evol 51:26–40

Zhang H, Hou Y, Miranda L, Campbell DA, Sturm NR, Gaasterland T, Lin S (2007a) Spliced leader RNA *trans*-splicing in dinoflagellates. Proc Natl Acad Sci USA 104:4618–4623

Zhang H, Bhattacharya D, Lin S (2007b) A three-gene dinoflagellate phylogeny suggests monophyly of prorocentrales and a basal position for *Amphidinium* and *Heterocapsa*. J Mol Evol 65:463–474

Zhang H, Bhattacharya D, Maranda L, Lin S (2008) Mitochondrial *cob* and *cox1* genes and editing of the corresponding mRNAs in *Dinophysis acuminata* from Narragansett Bay, with special reference to the phylogenetic position of the genus *Dinophysis*. Appl Environ Microbiol 74:1546–1554

Zhang Q, Liu G, Hu Z (2011) Morphological differences and molecular phylogeny of freshwater blooming species, *Peridiniopsis* spp. (Dinophyceae) from China. Eur J Protistol 47:149–160

Zhou C, Fernandez N, Chen H, You Y, Yan X (2011) Toxicological studies of *Karlodinium micrum* (Dinophyceae) isolated from East China Sea. Toxicon 57:9–18

Endosymbioses in Sacoglossan Seaslugs: Plastid-Bearing Animals that Keep Photosynthetic Organelles Without Borrowing Genes

Heike Wägele and William F. Martin

Abstract In this chapter, we summarize our knowledge on photosynthesis properties in the enigmatic gastropod group Sacoglossa. Members of this group are able to sequester chloroplasts from their food algae and store them for weeks and months in order to use them in a similar way as plants do.

Only four to five sacoglossan species are able to perform photosynthesis for months, others are less effective or are not able at all. The processes involved are not clear, but we show by this chapter that many factors contribute to the developing of a photosynthetic seaslug. These include extrinsic (environment, origin and properties of the nutrition and the plastids) and intrinsic factors (behaviour, physiological and anatomical properties). Maintenance of plastids is not enhanced by a horizontal gene transfer (HGT) from the algal genome into the slug genome, as was hypothesized for many years. We outline here the questions that now have to be asked and the research that has to be done to understand the factors that actually contribute to this unique metazoan phenomenon, which is not understood at all.

Contents

Introduction 292
History of the Famous Four 293
Investigation of Functionality of Plastids in Slugs from Different Perspectives 296
Ultrastructure 297
Photosynthetic Measurements 302
Physiology 304
Environmental Factors 305
Behaviour 306

H. Wägele (✉)
Zoologisches Forschungsmuseum Alexander Koenig, Adenauerallee 160, 53113 Bonn, Germany
e-mail: h.waegele@zfmk.de

W.F. Martin
Institut für Molekulare Evolution, Heinrich-Heine-Universität Düsseldorf, Universitätsstrasse 1, 40225 Düsseldorf, Germany

W. Löffelhardt (ed.), *Endosymbiosis*, DOI 10.1007/978-3-7091-1303-5_14,

© Springer-Verlag Wien 2014

Ecophysiology 307
Chloroplast Origin 307
Properties of Plastids 309
Functional Plastid Genomes and Lack of Horizontal Gene Transfer 309
Future Research 316
References 318

Introduction

"Leaves that crawl" is a term that Robert Trench (1975) aptly used to characterize the species of sacoglossan seaslugs that sequester plastids from their algal food source and maintain them in an active photosynthetic state, effectively allowing the slugs to grow on CO_2 and light. The slugs steal the plastids from the alga upon which they feed; hence, the plastids in the slugs are called kleptoplasts. When the plastids become dysfunctional they have to be replaced, the symbiosis has to be reestablished. But the kleptoplasts can remain photosynthetically active for several months in some species, a circumstance that has captured the imagination of many researchers (and journalists) over the last 50 years and that has time and time again fuelled discussion of the possibility of horizontal gene transfer from the genome of the algal food source into this fascinating metazoan life form (Rumpho et al. 2000, 2001, 2006, 2011; Wägele et al. 2010a; Pelletreau et al. 2011). The incorporation of functional kleptoplasts, referred to for simplicity as chloroplasts or plastids here, is not unique to sacoglossans, because many protists incorporate "previously owned" plastids into their metabolism (e.g. Rumpho et al. 2006; Lee 2006; Stoecker et al. 2009; Johnson 2011). The fascination concerning molluscan kleptoplasty from higher chlorophytes (usually Ulvophyceae) or a heterokontophyte (*Vaucheria litorea*, Agardh 1823) stems, however, from the circumstance that these fully fledged metazoans become photosynthetic for a good portion of their life cycle, such that associations with "little green men" of science fiction, and global nutrition questions of the "what if" type are never far. This chapter serves to first introduce the discovery of the famous four sacoglossan species that—instead of just eating their algal food like most of the about 300 known sacoglossan species—have independently evolved so as to undergo long-term retention of their plastids and then to summarize the literature dealing with the different research perspectives towards unravelling the mystery behind functional plastid retention. Finally, recent findings have opened up new insights into the genetic basis behind the functional relationship between the photosynthesizing algal organelle and its gastropod digestive gland system.

History of the Famous Four

Detection of chlorophyll in Sacoglossa goes back to the late nineteenth century (De Negri and De Negri 1876), but the surprising property of using an organelle, the "enslaved" chloroplast, for the slug's metabolic purposes was described by Kawaguti and Yamasu only in 1965, based on the electron microscopical studies of *Elysia atroviridis* Baba, 1955. These authors showed that the green structures housed in the digestive gland cells and surrounded by a double membrane were structurally identical to the chloroplasts in the ulvophycean *Codium fragile* (Hariot, 1889) upon which *E. atroviridis* feeds. They assumed that these chloroplasts perform photosynthesis, produce oxygen and, while within the digestive gland cells, "*. . .will give their products of photosynthesis to the host cell and will receive, in return, waste substances of the animal. . .*" (Kawaguti and Yamasu 1965, p. 60). Since then, many studies have been performed on sacoglossans to investigate this unique feature from various perspectives. This led to the discovery of sacoglossans that were much more effective in this symbiotic relationship now known as long-term retention (LTR) forms: *Elysia crispata* (Mørch, 1863) (Fig. 1d; Caribbean Sea), *Elysia timida* (Risso 1818) [Fig. 1b; Mediterranean Sea], *Elysia chlorotica* Gould, 1870 (Fig. 1e; North Atlantic) and *Plakobranchus ocellatus* van Hasselt 1824 (Fig. 1a; Indopacific). Trench et al. (1969) detected, and Clark and Bussacca (1978) described, the long-term retention of chloroplasts in *E.* (designated there as *Tridachia*) *crispata*. Greene (1970) described retention in *P. ocellatus* (described there as *P. ianthobapsus*), characterized in more depth recently by Evertsen et al. (2007). Rahat and Monselise (1979) and especially Marín and Ros (1989) described long-term retention in *E. timida*, and finally Graves et al. (1979) did so for *E. chlorotica*. Since these discoveries, many other sacoglossan species have been investigated and the plastid retention time (ranging from several days to several weeks) has been recorded (Händeler et al. 2009).

Of all species investigated so far, only *E. chlorotica* survives for more than 1 year of starvation (that is, survives from its plastids) in culture (Rumpho et al. 2000). Händeler et al. (2009) reported retention for nearly 3 months in *P. ocellatus*; however, unpublished data (H.W.) suggests retention possibly for 4 months and more. *E. timida* retains chloroplasts for 50 days (Wägele et al. 2010a) and *E. crispata* for more than 40 days (Händeler et al. 2009). A recently described *Elysia* species—*E. asbecki* (Wägele et al. 2010b) (Fig. 1e)—exhibits photosynthetic activity in the first 10 days, similar to the performance seen in *E. timida* and *E. crispata* (Wägele et al. 2010b). Therefore, it is possible that this species might be the second long-term retention form known from the Pacific. If so, that would increase the circle of sacoglossans with long-term plastid retention to five species.

Händeler et al. (2009) and Wägele et al. (2011) clearly showed that evolution of long-term retention occurred in separate lineages (Fig. 2), but the ability to incorporate plastids without direct digestion goes back to a common ancestor in the lineage of the Plakobranchoidea. This taxon is characterized by parapodia that can

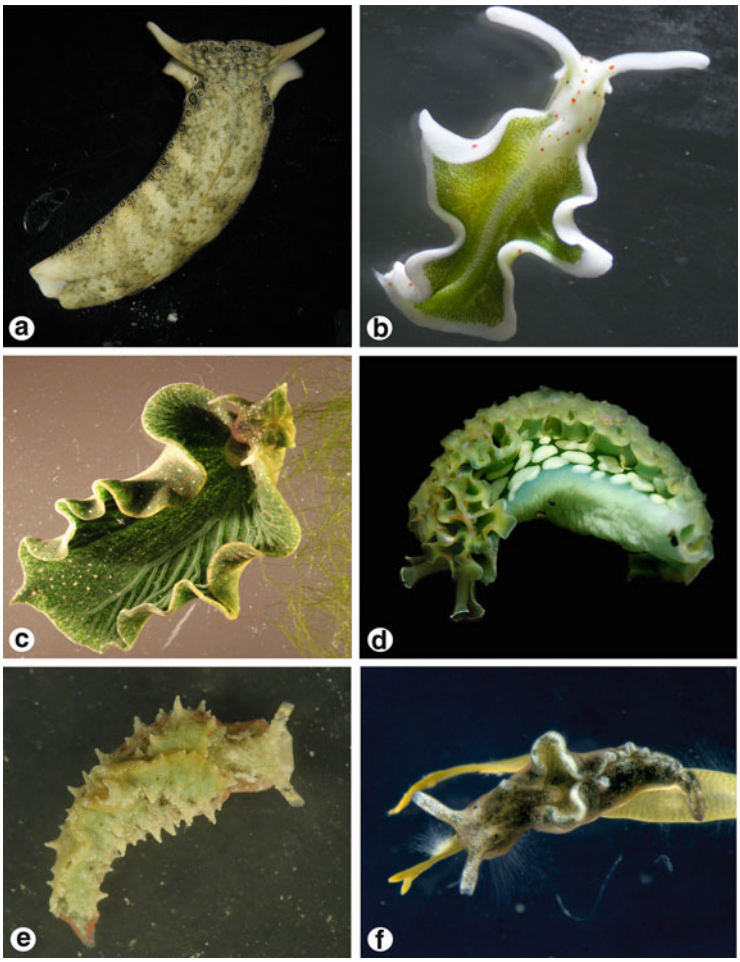


Fig. 1 The most intriguing sacoglossan species known for chloroplast retention. All species belong to the Plakobranchoidea, which typically have parapodia. (**a**) *Plakobranchus ocellatus*, (**b**) *Elysia timida*, (**c**) *Elysia chlorotica*, (**d**) *Elysia crispata*, (**e**) *Elysia asbecki* and (**f**) *Elysia viridis*. Photo of *E. chlorotica* provided with permission of M. Rumpho and K. Pelletreau

cover the whole dorsal body (Figs. 1a–f and 3d)—probably a key adaptation utilized by all long-term retaining species. Two other groups are described within Sacoglossa: the Limapontioidea that lack parapodia but can possess many flap-like dorsal appendages (Fig. 3a–c), and the shelled Oxynoacea (Fig. 3e, f). Maeda et al. (2010) contested the results of Händeler et al. (2009) by presenting an alternative phylogenetic hypothesis and claiming that kleptoplasty evolved at the base of the Sacoglossa, represented by the shelled *Cylindrobulla*. However, they used the term kleptoplasty to designate the act of merely engulfing and subsequently digesting chloroplasts. Immediate digestion is typical for nearly all

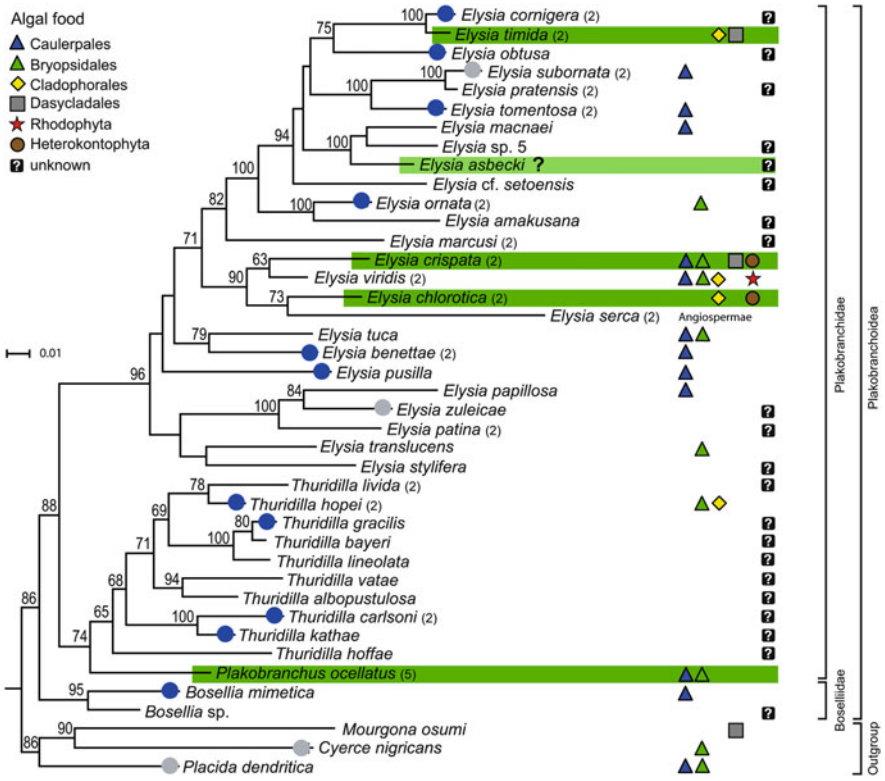


Fig. 2 Phylogeny of the Plakobranchoidea with information on food organisms [after Wägele et al. (2011)]. Maximum likelihood analysis of partial gene sequences (28S rDNA, 16S rDNA and *cox1*—first and second codon positions only). *Numbers* in *brackets* indicate the number of specimens included. *Numbers* above *branches* indicate bootstrap values. *Gray dots* and *blue dots* on the *branches* indicate non-retention or short-term retention of plastids, respectively. Long-term retention is pointed out by *green bars*, the fifth potential long-term form, *Elysia asbecki*, by a *light green bar* and a *question mark*

Limapontioidea and certainly all shelled Oxynoacea, it is clearly a trait ancestral within the Sacoglossa and should not be confused with kleptoplasty.

Clark et al. (1990) proposed six levels of chloroplast retention starting from no retention and direct digestion (level 1) to long-term function retention (level 6) with photosynthesis persisting for more than 1 week. Evertsen et al. (2007) extended this classification scheme by two additional levels in order to meet the capability of the long-term retention forms with a range of 1 week functional photosynthesis up to several months. With many species now known to exhibit a wide variety of chloroplast retention, and given the ambiguities of some reported photosynthetic measurements (due to environmental and individual factors, including the variety of ways to measure photosynthesis), these fine-grained levels of distinction are not used in the latest literature.

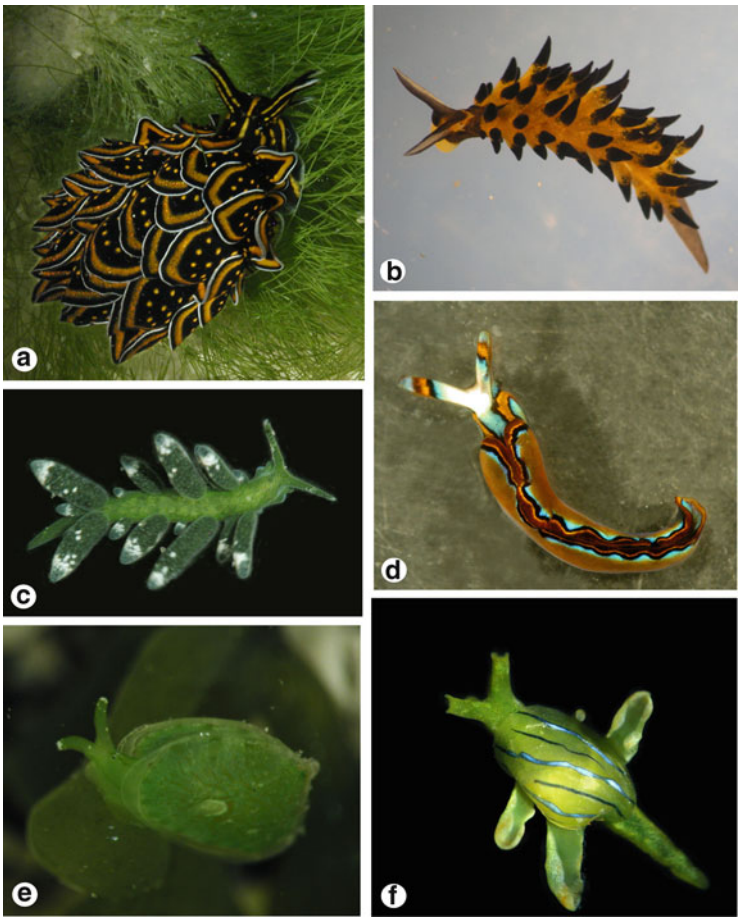


Fig. 3 Several sacoglossan members exhibiting only short-term or non-retention of plastids. (**a**–**c**) show members of the Limapontioidea, (**d**) of a Plakobranchoidea and (**e**, **f**) of shelled Oxynoacea. (**a**) *Cyerce nigricans*, (**b**) *Placida cremoniana*, (**c**) *Ercolania kencolesi*, (**d**) *Thuridilla neona*, (**e**) *Tamanovalva limax* and (**f**) *Lobiger viridis*

Investigation of Functionality of Plastids in Slugs from Different Perspectives

In the nearly 50 years following after the detection of functional plastids in the digestive gland cells of sacoglossans, the properties of the peculiar relationship between a plant organelle and a metazoan life form have been investigated from many perspectives. The following chapters review the various approaches.

Ultrastructure

The ultrastructure of sequestered plastids has been studied in several systems, usually in comparison with plastids in the natural algal source. Studies began with the investigation by Kawaguti and Yamasu (1965) on *E. atroviridis* and its food alga *Codium fragile*, followed by Taylor (1971) on *Hermaea bifida* (Montagu 1815) and its food *Griffithsia flosculosa* [now assigned to *Halurus flosculosa*, see Maggs and Hommersand (1993)]. Taylor (1968), Trench et al. (1973a, b) and Hawes (1979) studied *E. viridis* (Montagu, 1804) (Fig. 1f) feeding on C. *fragile*. Later Curtis et al. (2005, 2006) reported results for *E. crispata* and *E. clarki* (Pierce et al. 2006) and some of their food sources as *Bryopsis plumosa* (Hudson) (Agardh 1823), *Halimeda incrassata* (J. Ellis) Lamouroux, 1816 and *Penicillus capitatus* Lamarck, 1813. The ultrastructure of plastid retention was sometimes compared between long-term retention forms (LTR) and short-term (STR) or non-retention forms (NR): Graves et al. (1979) compared *E. chlorotica* (LTR) with a limapontioidean species, *Alderia modesta* (Loven, 1844) (NR); Evertsen and Johnsen (2009) compared *E. viridis* (STR) with *Placida dendritica* (Alder and Hancock, 1843) (NR).

These studies revealed a variety of plastid degradation stages within the digestive gland cells, depending on the species, its ability to incorporate plastids, as well as the duration of the starvation period. The occurrence of large starch grains in kleptoplasts and the higher frequency of plastoglobuli after several weeks of starvation—not observed in freshly fed species—is noteworthy [described, for example, for starved *Elysia viridis*, see Evertsen and Johnsen (2009)].

Of particular interest is the location of the plastid within the cytosol of the slug. Since plastids are incorporated via phagocytosis, one might expect three membranes to encircle an intact kleptoplast, one from the host phagosome as well as the inner and outer chloroplast membranes in the case of food plastids surrounded by two membranes; the (secondary) food plastids of *Vaucheria* are surrounded by four membranes in the alga. However, reports from the literature on this aspect from the various species remain incongruent (summarized in Table 1). *E. viridis* (STR), after 4 weeks starvation, has been reported to harbour a host membrane in addition to the intact plastid membranes. However, it was not observed in all sequestered plastids [Trench et al. (1973b), but see Hawes (1979) for different results] or the phagosome membrane appeared ruptured (Hawes and Cobb 1980; Evertsen and Johnsen 2009). Hirose (2005) investigated the ultrastructure of plastids in *P. ocellatus* (LTR), but came to no conclusion whether plastids were surrounded by host membranes or not. Muscatine et al. (1975) mentioned plastids with and without a phagosome membrane in *E. viridis* after feeding. Similar results were obtained for the limapontioidean species *Costasiella ocellifera* Simroth, 1895 (described as *C. lilianae*), which survived a 65-day starvation period after feeding on *Avrainvillea* (Clark et al. 1981). In this species, kleptoplasts were observed in the digestive gland cells with intact cp membranes, but not all were enclosed in a phagosome membrane. This was irrespective of starving condition

Table 1 Published results obtained from ultrastructural investigation of incorporated plastids with regard to presence of phagosome and chloroplast membranes

Species	PR	TEM results as outlined in the text	Number of membranes as specifically indicated in literature	Food and feeding	Authors
Alderia modesta	NR	Plastids surrounded by an "extrinsic" membrane, showing various degrees of degradation	3	*Vaucheria spec*	Graves et al. (1979)
Bosellia mimetica	STR	Plastids with a host membrane from a phagocytic vacuole	Probably 3	*Halimeda tuna*	Marín and Ros (1989)
Costasiella ocellifera (*as lilianae*)	?	Irrespective of plastid condition, a host membrane can be present or not. cp membrane is always intact		*Avrainvillea nigricans*	Clark et al. (1981)
Elysia australis	?	Plastids without host membrane			Marín and Ros (1993) [after Hinde (1983)]
Elysia chlorotica	LTR	Plastids bounded by an intrinsic double membrane and which contain parallel lamellae	2	*Vaucheria spec*	Graves et al. (1979)
Elysia chlorotica	LTR	Plastids with double membrane of cp and no host membranes	2		Rumpho et al. (2001)
Elysia clarki	LTR	Two membranes, the outer more loose, therefore probably from the host	2	*Penicillus spec.*	Curtis et al. (2006)
Elysia clarki	LTR	Plastids were separated from cytoplasm by a membrane		Various algae	Curtis et al. (2006)
Elysia maoria	?	Plastids without host membrane	2		Brandley (1981)
Elysia furvacauda	?	Plastids without host membrane	2		Marín and Ros (1993) Brandley (1984)
Elysia timida	LTR	Plastids with double chloroplast envelope; no host membrane	2		Marín and Ros (1989)

Elysia timida adults	LTR	Defunct plastids are surrounded by a host membrane. They are then digested	3		Marín and Ros (1993)
Elysia timida juveniles	LTR	Plastids with double chloroplast envelope and always a host membrane	3	*probably Cladophora*	Marín and Ros (1993)
Elysia translucens		Plastids with double chloroplast envelope; no host membrane!	2	*Udotea petiolata*	Marín and Ros (1989)
Elysia viridis	STR	Double membrane around plastids	2		Taylor (1968)
Elysia viridis	STR	Plastids envelope always intact, but external membrane outside the envelope sometimes present, sometimes not	2–3	*Codium fragile*	Trench et al. (1973b)
Elysia viridis	STR	Plastids surrounded by host membrane, but a similar number of plastids lack the animal membrane	2–3	After feeding	Muscatine et al. (1975)
Elysia viridis	STR	Plastids with host membrane and sometimes algal cytoplasm attached	3	*Codium fragile*	Hawes (1979)
Elysia viridis	STR	Plastids with host membrane after 28 days starvation. But membrane can be ruptured	3		Hawes and Cobb (1980)
Elysia viridis	STR	Feeding experiment: most plastids are intact with a phagosome membrane—these are the functional ones. Only few show a burst phagosome and double membrane with exposed thylakoid staples—they disintegrate	3 cp double membrane remains, phagosome and thylakoid membranes disintegrate	Fed with *Codium fragile*	Evertsen and Johnsen (2009)
Ercolania funerea	NR	Plastids with host membrane (Marín 1988)			Marín and Ros (1993) [after Marín (1988)]
Hermaea bifida	NR	Double membrane	2	*Griffithsia flosculosa*	Taylor (1971)

(continued)

Table 1 (continued)

Species	PR	TEM results as outlined in the text	Number of membranes as specifically indicated in literature	Food and feeding	Authors
Placida dendritica	NR	Feeding experiment: plastids intact with phagosome membrane and probably double membrane of cp. All degrade very quickly and simultaneously	3	*Fed with Codium fragile*	Evertsen and Johnsen (2009)
Placida dendritica	NR	10 min after feeding: Algal material in the lumen of digestive gland and in the vacuole containing the cp. Engulfing by microvilli Vacuole consisting of inner membrane of phagocytic vesicle	Probably 3	Starved for 4 days, then fed with *Codium fragile*	McLean (1976)
Thuridilla hopei	STR	Plastids with a host membrane from a phagocytic vacuole	3	*Cladophora vagabunda*	Marín and Ros (1989)

PR: period of retention as indicated in Händeler et al. (2009). *NR* non-retention, *STR* short-term retention, *LTR* long-term retention

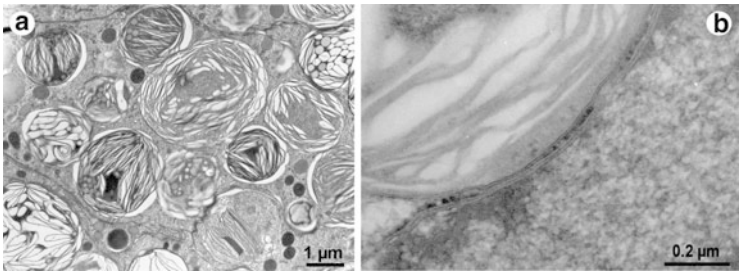


Fig. 4 Ultrastructural studies of incorporated chloroplasts in *Elysia timida* [after Wägele et al. (2011), originals with permission of R. Martin]. (**a**) *Acetabularia* chloroplasts in the cytosol of 2 months starved slug. (**b**) Chloroplast of fed slug in close contact to nuclear pore. Note the two membranes surrounding the chloroplast

and plastid condition. Marín and Ros (1989, p. 433) stated that the "*presence of a host membrane surrounding the chloroplasts in Thuridilla hopei and Bosellia mimetica seems to be closely related to fast plastid turnover*". That conclusion was supported by their results on two *Elysia* species, *E. timida* and *E. translucens* (Pruvot-Fol, 1957), which exhibited a much lower plastid turnover rate and showed no phagosome membrane.

The lack of a host membrane in *E. timida* but the presence of the two chloroplast membranes is confirmed in Wägele et al. (2011; Fig. 4). Interestingly, phagosome membranes exist in juvenile *E. timida* (Marín and Ros 1989). Similar to juvenile *E. chlorotica* [see Rumpho et al. (2000)], *E. timida* must first feed for a certain period of time before plastid retention with functional photosynthesis can develop. Brandley (1984) investigated *E. furvacauda* Burn, 1958, which not only feeds on *Codium* but also switches to other food organisms during the life cycle—including rhodophytes. They could not find a phagosome membrane in any of the incorporated plastids, although retention of the various plastid types was very different. *E. chlorotica* (LTR) show no phagosome membranes, but the plastids exhibit a double membrane, have direct contact to the cytosol and maintain their integrity even after 8 months of starvation (Mujer et al. 1996; Rumpho et al. 2001). Rumpho et al. (2001) emphasized the peculiar and unique situation in the relationship of *E. chlorotica* and its food *V. litorea*. Plastids in heterokontophyte alga *Vaucheria* exhibit four membranes: the double membrane of the plastid, followed by a third and fourth outer membrane, the latter of which is sometimes called the chloroplast endoplasmic reticulum. Ultrastructural investigations by Rumpho et al. (2001) confirmed the absence of the outer two membranes when the plastids are sequestered as kleptoplasts in the sacoglossan digestive gland cells. Rumpho et al. (2001) suggested that the outer two of the plastid's four membranes are probably also absent when plastids are transported in the lumen of the digestive gland.

In cases where a phagosome membrane is missing, plastids have direct contact to the cytosol of the slug's cell. Although the last word has not been spoken in regarding the presence or absence of a phagosome membrane in all these species,

the results to date suggest that functional kleptoplasts in long-term retention sacoglossans have a direct contact between the outer chloroplast membrane and the cytosol. Conversely, the presence of a phagosome membrane tends to indicate digestion. More investigations are needed, especially on STR forms such as *E. viridis*, regarding the number and nature of membranes surrounding the plastid at different stages of starvation. Clearly, a host phagosome membrane would serve to direct digestive enzymes towards the plastid, hardly a situation conducive to long-term retention. The situation in sacoglossans is perhaps similar to that observed for haptophyte kleptoplasts that are sequestered by the phagocytosing dinoflagellate *Dinophysis mitra*. Koike et al. (2005) reported various food particles and ingested cells within phagosomes, but never the ingested haptophyte kleptoplasts, leading them to conclude that the dinoflagellates "*. . .probably differentiate haptophytes from other prey, so that the haptophyte plastids are specifically treated not merely as food but as* "*special guests*" *in the D. mitra cytoplasm*" (Koike et al. 2005, p. 223).

The actual process of plastid engulfment has hardly been investigated so far but seems to occur in digestive gland cells quickly after feeding, according to McLean (1976), who studied the non-retention form *P. dendritica*. Evertsen and Johnsen (2009) described various stages of plastid degradation within *P. dendritica* (NR) and *E. viridis* (STR) after certain starvation periods. The latter species also show starch grains and nearly no recognizable remnants of plastids within phagosomes after several days of starvation.

Importantly, none of these ultrastructural investigations gave evidence of algal nuclei or nucleocytoplasm (Mujer et al. 1996) in the slugs' cells in any sacoglossan investigated yet. This is in contrast to some protists (Stoecker et al. 2009), where nuclei can be retained. Dividing chloroplasts were also never found in sacoglossans and plastids are not transmitted vertically via eggs (e.g. Trench et al. 1969; Marín and Ros 1993; Rumpho et al. 2001).

Photosynthetic Measurements

In early studies, various criteria for photosynthetic activity during starving periods were measured. Oxygen production was used by Graves et al. (1979) for *E. chlorotica* in comparison with the non-retaining form *A. modesta*. Taylor (1971) compared *H. bifida* with the food alga *G. flosculosa* and found a similar O_2 production during the first few days. This system is unusual so far in that *Hermaea* is not a member of the Plakobranchoidea, but of the Limapontioidea, which usually show no plastid retention at all, and the food alga is a rhodophyte. Chlorophyll contents were measured for several sacoglossans after various starvation periods (Clark and Bussacca 1978). Chlorophyll contents combined with the ability to fix labelled ^{14}C after various time periods were first used by Greene (1970) for *P. ocellatus* (LTR) and a short-term retention form, *E. hedgpethi* (Marcus 1961). Hinde and Smith (1972) used the same experimental design as

Greene (1970) and observed high photosynthetic activity in *E. viridis* over a period of 25 days. Marín and Ros (1989) and Clark et al. (1990) used similar methods for a survey of Mediterranean and Caribbean species. Carbon fixation, as a functional photosynthesis measurement, was further used by Hinde and Smith (1974) for some members of the Limapontioidea: *Limapontia capitata* (Müller 1774), *L. depressa* (Alder and Hancock 1862) and *A. modesta*. Their results on *Alderia*—exhibiting no photosynthesis—confirmed former studies [see also Graves et al. (1979)]. *L. capitata* (but not *L. depressa*) showed some carbon fixation during the first few days, but much less than *E. viridis*. The fixation rates reported in *C. ocellifera* by Clark et al. (1981) after 65 days of starvation corresponded to about 10 % of the algal rate, which appears very high since this species is a member of the Limapontioidea and not of the Plakobranchoidea. ^{14}C fixation rates measured in *E. chlorotica* (starved 4 months) were only slightly lower than in the food alga *Vaucheria* (Rumpho et al. 2001).

The incorporation of $^{14}CO_2$ as a proxy for photosynthetic activity in animals has a caveat, though. Animals can incorporate appreciable amounts of $^{14}CO_2$ into protein, lipid and carbohydrate (Agosin and Repetto 1963), but this can proceed through exchange reactions, without net CO_2 fixation (Louis Tielens and Manfred Grieshaber, personal communication). For example, many animals rely heavily upon propionate, a C_3 compound generated in the digestive tract, as the backbone of their carbon metabolism. The typical eukaryotic assimilation pathway for propionate involves carboxylation (incorporation of CO_2) to methylmalonyl-CoA, isomerization and thiolysis to succinate (Schöttler and Bennet 1991). In the presence of $^{14}CO_2$, this would lead to succinate (a C_4 compound) with label at either C-1 or C-4. En route to glucose, succinate is converted into triose phosphates (C_3 compounds) through decarboxylation, whereby the labelled or the unlabelled carboxyl group of succinate is removed with equal probability. This yields glucose with, on average, one labelled carbon atom for every two molecules of CO_2 incorporated, even though two carboxylations are balanced by two decarboxylations. The exchange reaction leads to incorporation of radioactive CO_2, but in a process that has nothing to do with photosynthesis.

In situ measurements of plastid function in Sacoglossa started with the use of a Pulse Amplitude Modulated (PAM) Fluorometer (Heinrich Walz, Germany) to detect in vivo fluorescence and to measure the quantum yield of charge separation in photosystem II in dark acclimated organisms to analyse the status of the photosynthetic activity (Wägele and Johnsen 2001). Further studies with PAM fluorometry on selected species followed (Evertsen 2006; Evertsen et al. 2007; Evertsen and Johnsen 2009; Jesus et al. 2010), including measurements of *P. ocellatus* showing nearly no decline of the quantum yield over a period of 2 months. Händeler et al. (2009) published measurements of nearly 30 species, including many genera never studied before. Yamamoto et al. (2009) focussed on Japanese sacoglossans, while Klochkova et al. (2010) focussed on sacoglossans in Korean waters.

Physiology

Physiological investigations were performed to understand the contribution of plastids towards energetic self-maintenance as well as supporting the slug. Trench and Smith (1970) detected synthesis of the photosynthetic pigments α- and β-carotene by tracing labelled ^{14}C in *E. viridis* and *E. crispata*; however, their results concerning xanthophyll pigments were negative and only small traces of chlorophyll *a* and *b* were detectable. Studies on *E. viridis* kept under starvation in light and dark conditions revealed a higher weight loss in darkness, implying that photosynthesis supports its metabolism (Hinde and Smith 1975).

Metabolites produced with the help of plastids were followed within slug tissues and traced to many precursor compounds, including lipids, amino acids, mono- and oligosaccharides (e.g. Hinde and Smith 1974). The role of different organs like neural tissues, the excretory system and the pedal gland have also been investigated (Trench 1969). Labelled ^{14}C was incorporated in the pedal gland of *P. ocellatus* (as *P. ianthobapsus*; Trench et al. 1973a) and *E. crispata* (as *Tridachia crispata* and *Tridachiella diomedea*; Trench 1973; Trench et al. 1974). Analysis of the mucus secretion after 3 weeks starvation revealed the presence of the photosynthetic products glucose and galactose, which are in turn precursors for the synthesis of mucus (Trench et al. 1969, 1970, 1973a, 1974). Subsequent studies confirmed these results for other species. Marín and Ros (1989) found about 7–8 % of net incorporated ^{14}C in the mucus of *E. timida*. Trench et al. (1973a, b) compared ^{14}C fixation in *E. viridis* and isolated plastids from their food (*C. fragile*) and showed that, while fixation rates are similar, the release of fixed carbon is much higher in the slugs than in isolated plastids. Also, the location of incorporated ^{14}C differs. When analysing *E. chlorotica*, Rumpho et al. (2001) mainly traced ^{14}C in water-soluble metabolites. In contrast, the food alga *Vaucheria* incorporated more ^{14}C in lipids and proteins.

Ireland and Scheuer (1979), analysing *P. ocellatus*, demonstrated a high incorporation of ^{14}C in secondary metabolites of the pyrone class (9,10-deoxytridachione). The in vivo photoconversion of this pyrone compound into photodeoxytridachione occurs when the animals are exposed to light. The same compounds have been found in the LTR forms *E. crispata* (Ireland and Scheuer 1979) and *E. timida* (Gavagnin et al. 1994). Ireland and Scheuer (1979) speculated that the photodeoxytridachiones function as a sunscreen, absorbing mainly ultraviolet light and therefore representing a biochemical adaptation to higher irradiances [but see Di Marzo et al. (1993), who assigned the function to defence]. Polypropionates seem to be more widespread within Sacoglossa, even in NR forms. The presence of these compounds within Plakobranchoidea, as well as Limapontioidea, was reported by Di Marzo et al. (1993) and Gavagnin et al. (2000).

Environmental Factors

Clark et al. (1981) showed the importance of environmental factors like temperature and irradiance on the survival of functional plastids within slugs. Net carbon fixation in *C. ocellifera* increased threefold from 20 to 30 °C, but declined again at 35 °C. The optimal temperature conditions for net carbon fixation in *E. timida*, measured in a range of 15–35 °C, were 25 °C (Marín and Ros 1989). Experiments with *E. tuca* Marcus and Marcus, 1967 in situ at its main locality (along the Key West, Florida) clearly showed seasonal and geographical variation in chlorophyll content (Waugh and Clark 1986, p. 485). These authors indicated that many factors, "*including feeding rate*, *metabolic rate of plastids*, *physiological support of kleptoplastids*, *and dietary selection may affect chlorophyll level*"—and they found the highest chlorophyll content during autumn period, when the animals also show highest reproductive activities.

High light intensity elicited few immediate effects on the direct carbon fixation in *C. ocellifera* (Clark et al. 1981), but observations over longer periods clearly indicate photodamage and subsequently lower carbon fixation compared to animals kept in lower irradiances (Clark et al. 1981). Marín and Ros (1989) found the highest carbon fixation in *E. timida* under light intensities of 200 μE m^{-2} s^{-1}. Giménez Casalduero and Muniain (2008) investigated chlorophyll content and oxygen production in starved *E. timida* kept in a normal day/night rhythm or kept in permanent darkness. Chlorophyll content decreased in similar amounts during the 28 days of starvation, irrespective of exposure to 12 h irradiance with 200 μE m^{-2} s^{-1} (for comparison, cloudless noontime irradiance in the tropics is about 1700 μmol photons/m^{-2} s^{-1}). However, the animals kept in light showed less shrinkage than those kept in the dark. This probably confirms the mentioned irradiance as an optimum for *E. timida* with low photodamage to the photosystems. Vieira et al. (2009) demonstrated that maximum quantum yield values decreased considerably quicker when specimens of *E. viridis* were exposed to medium light intensities (140 μE m^{-2} s^{-1}) during starvation experiments over a maximum of about 8 days, compared to lower light intensities (30 μE m^{-2} s^{-1}). Based on several analyses, including rapid light response curves (RLC), they suggested that photodamage of the D1 protein in the PSII reaction centre and the absence of repair mechanisms is the major factor of this decline. Evertsen and Johnsen (2009) reported similar results concerning the same species. Photoacclimation is a typical feature of plants adapting to changing light conditions, expressed by the angle α of the E_K curves when rapid light responses are measured. Vieira et al. (2009) could not detect any changes in α during low and high light experiments, inferring that no acclimation had occurred during the experiments. The chloroplasts showed the acclimation status of the light regime, when the slug consumed the alga. *E. chlorotica*, which incorporates plastids from the heterokontophyte *V. litorea*, was found to maintain functional levels of thylakoid membrane proteins over several months [see Mujer et al. (1996), Pierce et al. (1996), Green et al. (2000)]. *E. chlorotica* is the only species where evidence is given for active protein synthesis

in kleptoplasts (see below). However, studies on isolated *Acetabularia* plastids—the food of *E. timida*—indicate that the synthesis of chloroplast pigments (α- and β-carotene, chlorophyll *a*, chlorophyll *b* and xanthophylls) is still possible (Trench and Smith 1970) and therefore more studies on other long-term retention forms are needed.

Behaviour

Behavioural studies are few in number and were mainly performed on the Mediterranean species *E. timida*. Their major target was the importance of parapodia and their role in shading, therefore reducing photodamage in chloroplasts. Rahat and Monselise (1979) and Monselise and Rahat (1980) investigated the importance of parapodial movements in *E. timida*, first in the laboratory and subsequently in the sea, resulting in clear evidence of the parapodia closing over the body in darkness and very high irradiances. They showed that the eyes of the slugs are not important for measuring the light intensity, since opening degrees of the parapodia were similar in normal or eyeless slugs in various light regimes. Giménez Casalduero and Muniain (2008) interpreted the 65 % loss in chlorophyll content, but only a reduction of 25 % of photosynthetic production rate after 9 days starvation in *E. timida* as evidence for the ability of the slug to regulate its exposure to incident light by unfolding and refolding of its parapodia. Schmitt and Wägele (2011) reported results in support of that view by measuring fluorescence values in *E. timida* during various folding positions of the parapodia. This slug species clearly showed a strong, positive, phototactic behaviour, whereas reactions of *T. hopei* (Verany 1853), a close relative, were distinct, the latter species preferring dim light conditions. Since juvenile specimens of *E. timida* showed a strong phototactic reaction already before their first feeding, Schmitt and Wägele (2011) suggested that this behaviour might not be triggered by the chloroplasts.

Only one study addresses the recognition of light spectra. Weaver and Clark (1981) tested one long-term retention form (*E. crispata*), two short-term forms (*E. tuca* and *C. ocellifera*, as *C. lilianae*) and two non-retention forms [*Oxynoe antillarum* Mørch, 1863 and *Berthelinia caribbea* Edmunds, 1963]. Animals were allowed to choose between two different light spectra in two-trail experiments. Cellulose acetate filters were used with transmission of red, blue, yellow or green light, but wavelengths were not given. According to their interpretation, the three photosynthetic animals preferred green, yellow and blue light, thus shorter wavelengths, which at the same time correlated with higher light intensities. The two non-retention forms showed no light quality preferences but in general thrived under low light intensities or even darkness.

Ecophysiology

Jesus et al. (2010) addressed an ecophysiological problem not considered before in slugs: non-photochemical quenching (NPQ), typical for plants when quenching excessive light energy and removing excess energy from the photosynthetic electron transport chain by using accessory pigments. In *Acetabularia acetabulum* (Linnaeus) Silva, 1952 and *E. timida*, they clearly showed that pigments associated with the xanthophyll cycle behaved similarly in various light regimes and "were linearly related to NPQ capacity" (Jesus et al. 2010, p. 103). This was the first evidence that photo-regulation via the xanthophyll cycle, one of the most important photo-protection mechanisms in plants, can also operate in kleptoplasts isolated within slugs. They concluded that, in combination with behavioural responses, this mechanism could be important in retaining functional chloroplasts. Unusually however, a high xanthophyll ratio (higher zeaxanthin values) was observed in the dark experiments with *E. timida*. This was very different from *Acetabularia* kept under the same conditions. Since plastids in *E. timida* are not located within a phagosome membrane, resulting in a direct contact with the slug's cytosol, Jesus et al. (2010) suggested that there might exist some kind of interaction between the animal's metabolism and the electron transport chain of the kleptoplast. Another problem was addressed by Teugels et al. (2008), namely how the slugs overcome nitrogen deficiency during food shortages. By tracing ^{15}N-labelled ammonium, urea, nitrite and nitrate in dark and light experiments using *E. viridis*, they showed that N-uptake was higher during light exposure. However, nitrate was not used. They also measured ^{15}N uptake in dark green animals containing a high chloroplast density compared to light green animals with lower chloroplast density. When exposed to light, the uptake of ^{15}N molecules was much higher in the darker than in the lighter animals. Evidence of assimilation was observed by the presence of amino acids with ^{15}N signatures.

Chloroplast Origin

Early on, the species of food ingested, hence the type of chloroplasts sequestered, was considered as an important factor for functional photosynthesis in Sacoglossa. Accordingly, food algae were investigated by various authors with regard to performance in slugs' photosynthesis (Clark and Bussacca 1978). Food information was usually obtained by observing the slugs on specific substrates or during active feeding (e.g. Jensen 1980, 1993a, b; Thompson and Jarman 1989, for the most recent compilation see Händeler and Wägele 2007). Occasionally, pigment analyses were utilized to determine the origin of the incorporated plastids. For example, Greene (1970) rejected the former identification of cyanobacteria in *P. ocellatus* by Kawaguti and Yamasu (1965) based on the presence of chlorophyll *b*, which is not present in heterokontophytes but in chlorophytes. Additionally, the

combination of various pigments (especially siphonein and siphonaxanthin) indicated chlorophytes as the probable plastid source in *E. crispata* (Trench et al. 1969). Jesus et al. (2010) analysed pigment composition in *E. timida* and identified *Acetabularia* as the only probable food source. The conclusion that *H. filicina* is the food source for *E. timida* based on the observation that the slug was seen sitting on the alga [see Giménez Casalduero and Muniain (2008)] is tenuous.

Recent studies focus on barcoding the chloroplasts using single gene analyses (mainly *rbcL* and *tufA*) (Curtis et al. 2006; Händeler et al. 2010) to identify the source of ingested plastids. This is especially important when species are never observed to be feeding on algae (for example *P. ocellatus*), or when they feed on a variety of algae (for example *E. crispata* and *E. clarki*). Pierce et al. (2006) identified at least four different food items in *E. crispata*. A similarly diverse diet was confirmed for *E. clarki* (Curtis et al. 2006). Wägele et al. (2011) identified several food items belonging to different algal genera for the long-term retention form *P. ocellatus* by analysing *tufA*. From feeding and starvation experiments (Klochkova et al. 2010), it seems very unlikely that different kinds of sequestered plastids contribute in the same way and remain for the same time period. Therefore, the identification of retained plastids after weeks and months, especially in *E. crispata*, *E. clarki* and *P. ocellatus*, is essential to understanding the system. Interestingly, STR and NR forms can feed on the same algal species [*V. litorea*, see Pierce et al. (1996); *Caulerpa racemosa* (Forsskål) Agardh, 1873, see Händeler et al. (2010); *Chaetomorpha* spp., *Codium* spp., see Klochkova et al. (2010)] implying that long-term retention forms actively avoid direct chloroplast digestion.

A few studies deal with the impact of the alga on the feeding process. Waugh and Clark (1986) observed that the chlorophyll content varied in *Elysia tuca* according to the food organism on which the animals were collected. Animals collected from *H. incrassata* showed higher chlorophyll content than animals collected from *Halimeda discoidea* (Decaisne 1842). The authors considered the effort a slug must expend sucking the cytoplasm from the alga an important factor. Both *Halimeda* species calcify differently and therefore sucking might be more difficult in the more calcified *H. discoidea*. The relationship between *A. acetabulum* and *E. timida* has been well studied because of the close interaction during their annual life cycles: Juvenile slugs are able to accumulate chloroplasts, but grazing decreases when the alga increases calcification. The slug reaches its peak plastid accumulation when the alga is resistant to further grazing. Marín and Ros (1992, 1993) postulated the disappearance of food sources as a driving force for development of plastid retention in *E. timida*. Waugh and Clark (1986) hypothesised that algal wound-plug formation might retard the feeding process and therefore limit the uptake of chloroplasts. Jensen (1994) suggested algal cell wall structure, chemical composition and calcification as the most important factors governing the evolution of food preferences within Sacoglossa.

Händeler et al. (2009) mapped available food information on a phylogenetic tree of the Sacoglossa (Fig. 2, complemented with recent data). Diet data is however still unknown for many species. No pattern can currently be seen that would indicate a

preferred group of algae in the evolutionary process leading up to solar-powered sea slugs. We only know for sure that the algae have to be siphonaceous or siphonal (coenocytic), because sacoglossans suck out the contents of the filamentous algae and do not graze like members of the Anaspidea.

Properties of Plastids

The longevity and properties of isolated plastids as an indication of suitability for long-term retention was investigated in few cases. Giles and Sarafis (1972) isolated plastids from *Caulerpa sedoides* and maintained them in a functional state for more than 10 days within hens' eggs. CO_2 fixation was reduced to 48 % compared to the intact alga and decreased to 10 % after 27 days. Trench et al. (1973a) analysed carbon fixation rates in isolated *C. fragile* plastids. Fixation of ^{14}C still occurred after 5 days and is thus similar to results obtained from *A. acetabulum*, the exclusive food of *E. timida* [see Trench et al. (1973a, b)]. The authors discussed the putative robustness of siphonaceous algal plastids as an important factor in surviving engulfment into the slug's digestive cells. The stability of isolated plastids from *Codium* and *Caulerpa* was already emphasized by Grant and Borowitzka (1984a, b). Isolated *Vaucheria* plastids also exhibited a high structural and functional stability (Rumpho et al. 2001; Green et al. 2005).

Grant and Borowitzka (1984a) stressed the possibility that a contamination by extraplastidic material, enclosed in the so-called cytoplast, could enable the isolated plastids to produce many products (sugars and amino acids) for a long time. However, there are differences in the photosynthetic products when comparing isolated plastids from *Caulerpa*, *Acetabularia*, *Codium* and *Bryopsis*. Evidence for the release of photosynthetically produced substances from isolated plastids into the surrounding medium is also shown, but the uptake of exogenous material by isolated plastids seems to be very low [see Grant and Borowitzka (1984b)].

The autonomy of plastids as an important factor was also addressed by Rumpho et al. (2001). They assumed a higher autonomy of *Vaucheria* plastids due to the larger plastid genome, which contains many more genes necessary for photosynthesis, as compared to chlorophyte plastid genomes. But many LTR species retain chlorophyte plastids, so plastid gene content can hardly be a direct factor.

Functional Plastid Genomes and Lack of Horizontal Gene Transfer

How can a single organelle with a limited number of genes—between 60 and 200 protein coding genes for plastids (Timmis et al. 2004)—function without the complementary nuclear DNA that encodes for many proteins and enzymes necessary to uphold a functional photosynthesis over a period of several months? Less

than 10 % of the roughly 200–400 proteins required for photosynthesis in plastids are encoded by the plastid genome (Martin and Herrmann 1998) and thousands of genes that were acquired from cyanobacteria are present in the higher plant nuclear genome (Martin et al. 2002). That genes were transferred to the host nucleus during the origin of plastids from cyanobacteria has been known for some time (Martin and Cerff 1986; Martin et al. 1993, 1998); hence, the possibility that genes were also transferred during kleptoplast establishment has been considered as well.

The presence in sacoglossans of photosynthetically active sequestered plastids in the absence of sequestered algal nuclei led Pierce et al. (1996) to suggest that a horizontal gene transfer might have occurred from the algal nuclear genome into the slugs' nuclear genome. This would then enable the slug to encode the necessary proteins, subsequently targeted to the plastids they have sequestered. Their analyses with western blots on isolated plastids from *E. chlorotica* in comparison with plastids from the food alga *V. litorea* revealed a similar spectrum of proteins in both organisms. RuBisCo synthesis was experimentally interrupted in the slug by adding chloramphenicol, indicating plastid gene encoding. The authors concluded, however, that plastid proteins had to be synthesized on slugs' ribosomes within the slug cells, although, they did not mention the number of days the slugs were starved before performing the experiments. Pierce et al. (1996) discussed three possibilities as to where the genetic information could have come from. The first hypothesis suggests that the slug genome is genetically enriched via horizontal gene transfer (HGT) from the algal genome. The second proposes that the slug genome itself is already capable of encoding proteins targeted to the plastids. This is mentioned as dual targeting in Rumpho et al. (2000): The similarity of some proteins used in mitochondria and in plastids might simplify the process of redirecting. Rumpho et al. (2001, p. 310) considered that dual targeting "*. . .opens the possibility for animal proteins to be 'mis-directed' to the kleptoplasts*". The third hypothesis suggests that plastids release RNA to the slug cytosol, where the necessary proteins are synthesized and subsequently transported back to the plastids. Mujer et al. (1996) mentioned a fourth possibility, that kleptoplast proteins are very stable.

Mujer et al. (1996) stressed the necessity of molecular analyses for understanding the extraordinarily long plastid retention in the slugs. They reported the identification of several plastid-derived genes in isolated DNA from starved *E. chlorotica* using Southern blot analyses (*psaB*, *psbA*, *rrn16*, *rbcS* and *rbcL*). Protein products (D1, D2 and CP43) were still detected after several months of starvation. "*The ability of the symbiotic plastids to carry out transcriptional and translational functions helps explain their capacity for maintaining long-term photosynthetic activity*" (Mujer et al. 1996, p. 12336). The combination of functional proteins from PSI and PSII in the first few months of starvation, and the presence of certain larger proteins in an 8-month starved specimen, assigned to be products of *psaA* or *psaB*, led to their conclusion that PSI probably lasts longer in later starvation stages than PSII. Analyses by Green et al. (2000) suggested that kleptoplasts in *E. chlorotica* maintain photosynthetic oxygen production and electron transport activity through PSI and PSII for at least 5 and 6 months, respectively. Several proteins were identified by immunoblotting after even 7–9 months of

starvation, although some of them decreased considerably within this time. The identified proteins included some from PSII (D1, D2 and CP-43), PSI (PsaA, PsaB, PsaC and PsaD), and the cytochrome b_6/f complex (cytb_6 and cyt$f$), mobile electron carriers (cytc_6), RbcL, and the light-harvesting complex (FCP). Of these, only the last one is nuclear encoded. Translational ability was investigated by exposure to ^{35}S-labelled methionine in starved slugs and in the algae. Synthesis of D1, RbcL and several unidentified proteins was demonstrated even after 9 months, however not in the case of FCP. The authors discussed an extreme stability of this protein or a possible HGT for the specific gene. In general, the results confirmed former experiments (Mujer et al. 1996) that PSI remains functional longer than PSII. Due to negative results concerning presence of the algal ITS region in *E. chlorotica*, Green et al. (2000) ruled out the presence of an algal nucleus in the tissue of the slugs.

Rumpho et al. (2000) considered specifically the gene for phosphoribulokinase (PRK) as a possible candidate for HGT, since the encoded enzyme is needed for regenerating the CO_2 acceptor ribulose-1,5-bisphosphate to uphold photosynthetic CO_2 reduction (the Calvin cycle). In 2001, the authors reported PRK staying active for 9 months, but Southern blots and reverse transcriptase-PCR analyses were negative concerning the genes encoding PRK, RbcS and FCP. Rumpho et al. (2001, p. 310) did not entirely reject HGT as a consequence, but emphasized again the possibility that incorporated chloroplasts could be "*incredibly stable*". Hanten and Pierce (2001) published results on proteins belonging to the light-harvesting complex I (LhcI) in *E. chlorotica*. Since protein accumulation was blocked by cycloheximide, the authors inferred a nuclear-encoded synthesis from transferred genes in the slugs. The pigment protein FCP was again analysed by Pierce et al. in 2003, but they used another LTR species: *E. crispata*. Western blot analyses resulted in similar observations to those seen in *E. chlorotica*; cycloheximide appeared to block the synthesis of the protein. Pierce et al. (2003) tried to verify the identification of this protein through purification. The following determination of the N-terminal amino acid sequence (30 AA) and a subsequent BLAST search revealed 66 % sequence identity with FCP from the chromophyte *Cylindrotheca fusiformis*. The identification of three internal sequences lead to an 81 % overlap of one internal sequence (11 amino acids) with FCP protein sequence of the chromophyte *Macrocystis pyrifera*. Subsequently, the authors amplified a sequence (350 bp) of *Vaucheria* DNA after designing degenerate primers. Instead of applying the same primers to retrieve a sequence directly from the slugs, they indirectly identified part of the sequence by using Southern blot analyses against DNA taken from *E. crispata*. The resulting protein sequence showed a 51 % match with the FCP seen in *Laminaria saccharina*. These results, combined with western and Southern blot analyses, led the authors to infer that they had identified a gene that encodes for FCP within "*the genomic DNA of E. crispata, where it waits for the acquisition of new plastids in each generation of slugs*" (Pierce et al. 2003, p. 239). Pierce et al. (2007) reported evidence for gene sequencing of a transferred and targeted protein. The authors designed primers for FCP, lhc 1 and 2 (designated as lhcv 1 and 2 and as identified in former studies). The process of designing the

primers for the genes is not outlined, but subsequent sequencing of every amplicon of the three different genes revealed 100 % matches with respective genes in *Vaucheria* and *E. chlorotica*. They did not discuss however, whether these sequences match or mismatch respective GenBank sequences. They concluded that, due to the many unidentified proteins apparently blocked by cycloheximide and detected in former analyses, the number of genes transferred from the algal nucleus into the slug's nucleus must be very high.

Rumpho et al. (2008) then published results suggesting the presence of the algal gene *psbO* in *V. litorea* as well as in *E. chlorotica*. *PsbO* encodes MSP (the manganese stabilizing protein), a major protein associated with the oxygen evolving complex of photosystem II. They first designed a specific primer based on available *psbO* sequences in the GenBank database. Its product only showed a 48–68 % match to several MSP amino acid sequences from red algae. Nevertheless, by using the RACE method, they created a larger *psbO* product from *Vaucheria* nucleic acid extracts containing nearly 1,000 bp. Subsequently, new primers were designed from this long sequence. Amplification with these primers yielded PCR products with a 100 % similarity in comparisons of products obtained from *Vaucheria* and *E. chlorotica* nucleic acid extracts. A distressing aspect of that study is the description in detail of a bipartite targeting sequence in the putative *psbO* gene of the slug that would direct the product to the plastid. The plastids of *Vaucheria* are indeed surrounded by four membranes in the alga (Gould et al. 2008) and do require such complex targeting signals to traverse the four membranes, but the *Vaucheria* plastids sequestered in *Elysia* are only surrounded by two membranes, such that if those targeting signals were indeed present, the protein would exit the slug cell via the secretory pathway, and not be targeted to a plastid.

Rumpho et al. (2009) announced the presence of at least parts of the PRK gene (identical with *V. litorea*) in *E. chlorotica*. Activity of PRK protein was demonstrated for starved animals after 3 months, but redox regulation occurred only in PRK when studied in *Vaucheria* and not in the slug. Subsequent further identification of genes putatively subject to HGT was based on an EST analysis of *V. litorea* (Pierce et al. 2009; Schwartz et al. 2010). Knowing which genes are nuclear encoded in *Vaucheria* and necessary for upholding photosynthesis, they searched the *Vaucheria* EST library and subsequently designed primers. This resulted in algal sequences that exactly matched the slug's and are assigned to uroD, chlD, chlH, chlG, lhcv-3, lhcv-4 and PRK (Pierce et al. 2009; Schwartz et al. 2010). Although UroD (uroporphyrinogen decarboxylase) is synthesized in mitochondria and cytoplasm in animals, the authors rejected a possible dual targeting, at least for this gene. They suggested that the slug's UroD, which had only 27 % amino acid identity with UroD of the alga, is not responsible for the ongoing photosynthesis in *E. chlorotica*.

Notably, some of the identified genes (PRK, lhcv-3 and chlG) have an intron in *Vaucheria*, which is missing in the corresponding sequence obtained from *E. chlorotica*. Pierce et al. (2009) suggested that due to the high similarity of the genes in *Vaucheria* and *E. chlorotica*, the HGT must have been a very recent evolutionary event. They speculate that many more algal genes have been

transferred, "...—*perhaps even pieces of, or even entire, algal chromosomes are involved*" (Pierce et al. 2009, p. 127). Concerning the Sacoglossa in general, they also suggested that "...*gene movements have occurred many times across species and in different amounts*" (Pierce et al. 2009, p. 127).

In contrast, Rumpho et al. (2011, p. 307) emphasized that nearly all of the detected enzymes (except of two) are also encoded by the nuclear genome of animals. "...*it is possible that the animal could provide substitute proteins for the majority of the nuclear-encoded RPPP* (Calvin cycle) *enzymes if they were properly targeted to the foreign plastids*". This is possible, in principle at least, because the Calvin cycle and glycolysis/gluconeogenesis share a number of enzyme activities in common, and many Calvin cycle enzymes are indeed evolutionarily derived from duplicates of genes for cytosolic proteins that existed prior to endosymbiosis in the host cell that acquired the cyanobacterial ancestor of plastids (Martin and Schnarrenberger 1997).

Thus, in about 1996, the notion of HGT in sacoglossan kleptoplasty started to snowball, and it turned into a small avalanche with the 2008 report of the "transferred" psbO gene, the nuclear localization of which in the slug was not unequivocally shown, and which was obtained through PCR, not through a direct clone library (Rumpho et al. 2008). From the standpoint of classical endosymbiotic theory, one problem stood out in the *Elysia* gene transfer story that made us especially critical of the HGT claims. One of the crucial lines of reasoning behind the idea that plastids are derived from endosymbiotic cyanobacteria to begin with was the continuity of plastids through the egg cells of each generation, as Schimper (1883) and Mereschkowsky (1905) argued over 100 years ago. In other words, if the genes to support the plastids are present in sacoglossan slug nuclei, as some are claiming, why do the slugs reacquire the plastids every generation?

We reasoned that if there are transferred genes for photosynthetic functions in the nuclei of LTR slugs, we should be able to see those genes as expressed mRNA in a deep sequencing EST experiment using mRNA extracted from photosynthesizing slugs. We performed EST analyses of two LTR species, *P. ocellatus* and *E. timida* (Wägele et al. 2011) with 77,000 expressed sequence contigs for *P. ocellatus* and 25,000 contigs for *E. timida* (a total of 1.5 million reads and 64 Mb of nonredundant sequence data), made against mRNA that was extracted from animals that were demonstrably photosynthetic at the time of harvesting (PAM fluorescence) but removed from their food source for at least 3 weeks. We then compared the extensive *Arabidopsis* EST data, where a wealth of information on nuclear encoded genes for chloroplast biogenesis exists, to our slug ESTs and to the limited EST data then available for *Acetabularia*, the food alga of *E. timida*. The comparison to *Acetabularia* tells us whether, using *Arabidopsis* query sequences, we would be able to detect expressed algal genes in *Elysia* if they were there (a positive control for our computer analyses) while the comparison to the slugs tells us which homologues of nuclear encoded *Arabidopsis* genes for chloroplast proteins are expressed as mRNAs in the animals. The results (Fig. 5) clearly rule out the expression, by slugs, of horizontally transferred genes, from algae, as a component of plastid survival and functional photosynthesis in these two (out of the four

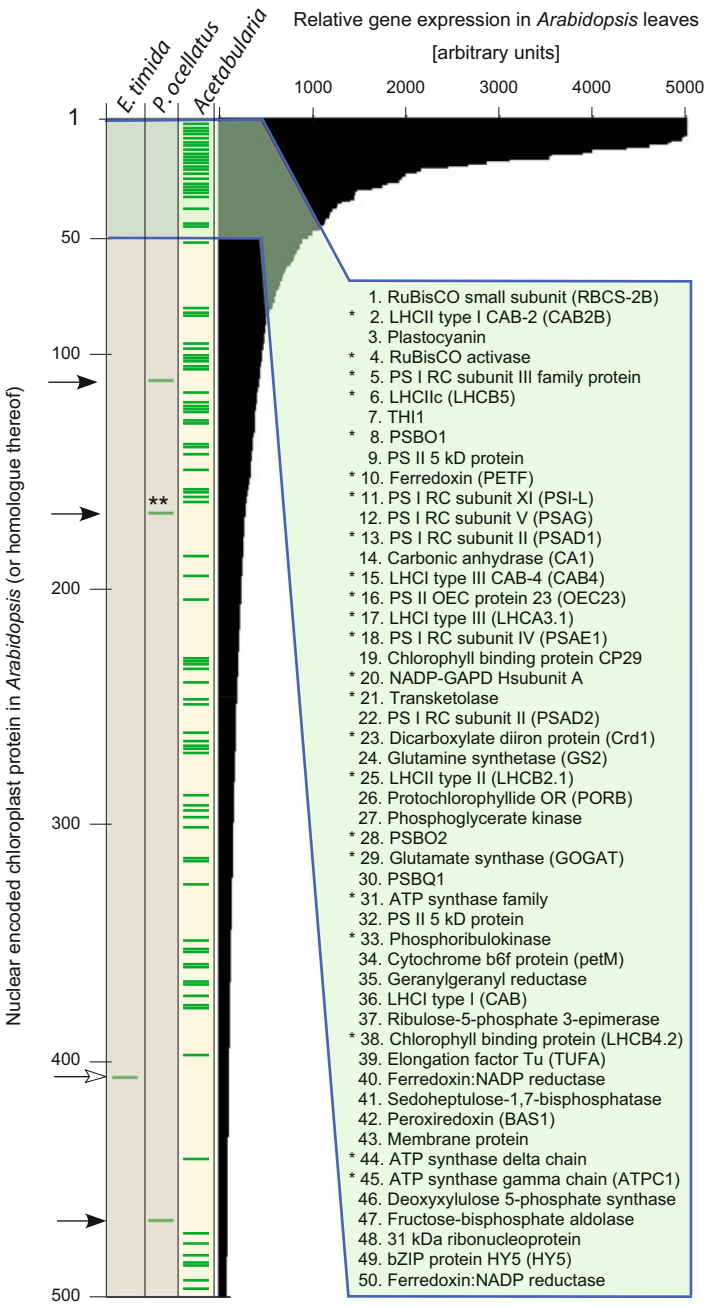

E. timida
P. ocellatus
Acetabularia
Relative gene expression in Arabidopsis leaves
[arbitrary units]
1000
2000
3000
4000
5000
1
50
100
200
300
400
500
Nuclear encoded chloroplast protein in Arabidopsis (or homologue thereof)
**
1. RuBisCO small subunit (RBCS-2B)
* 2. LHCII type I CAB-2 (CAB2B)
3. Plastocyanin
* 4. RuBisCO activase
* 5. PS I RC subunit III family protein
* 6. LHCIIc (LHCB5)
7. THI1
* 8. PSBO1
9. PS II 5 kD protein
* 10. Ferredoxin (PETF)
* 11. PS I RC subunit XI (PSI-L)
12. PS I RC subunit V (PSAG)
* 13. PS I RC subunit II (PSAD1)
14. Carbonic anhydrase (CA1)
* 15. LHCI type III CAB-4 (CAB4)
* 16. PS II OEC protein 23 (OEC23)
* 17. LHCI type III (LHCA3.1)
* 18. PS I RC subunit IV (PSAE1)
19. Chlorophyll binding protein CP29
* 20. NADP-GAPD Hsubunit A
* 21. Transketolase
22. PS I RC subunit II (PSAD2)
* 23. Dicarboxylate diiron protein (Crd1)
24. Glutamine synthetase (GS2)
* 25. LHCII type II (LHCB2.1)
26. Protochlorophyllide OR (PORB)
27. Phosphoglycerate kinase
* 28. PSBO2
* 29. Glutamate synthase (GOGAT)
30. PSBQ1
* 31. ATP synthase family
32. PS II 5 kD protein
* 33. Phosphoribulokinase
34. Cytochrome b6f protein (petM)
35. Geranylgeranyl reductase
36. LHCI type I (CAB)
37. Ribulose-5-phosphate 3-epimerase
* 38. Chlorophyll binding protein (LHCB4.2)
39. Elongation factor Tu (TUFA)
40. Ferredoxin:NADP reductase
41. Sedoheptulose-1,7-bisphosphatase
42. Peroxiredoxin (BAS1)
43. Membrane protein
* 44. ATP synthase delta chain
* 45. ATP synthase gamma chain (ATPC1)
46. Deoxyxylulose 5-phosphate synthase
47. Fructose-bisphosphate aldolase
48. 31 kDa ribonucleoprotein
49. bZIP protein HY5 (HY5)
50. Ferredoxin:NADP reductase

Fig. 5 Expressed genes in *Elysia timida* and *Plakobranchus ocellatus* compared to highly expressed homologous nuclear genes mainly encoding for photosynthesis proteins from the food alga of *E. timida*, *Acetabularia acetabulum*, and from *Arabidopsis* [after Wägele et al. (2011)]. For details see original publication. The most highly expressed 50 genes in *Arabidopsis* are listed. Matches with the *Acetabularia* contigs is indicated with an *asterisk*. Slugs' contigs showed no

well known) LTR sacoglossan species (Wägele et al. 2011). These results mean that the longevity of the sequestered plastids does not require transferred genes. The slugs have no need for HGT to support their photosynthetic symbiosis. We conclude that they just sequester long-lived plastids—which is possibly even more interesting.

But the HGT story is hard to stop, even with strong data for expressed genes of the type we presented. Pelletreau et al. (2011) repeated our EST experiment and analysed *E. chlorotica*, reporting 148 Mb of sequence data. They also found no evidence for expression of transferred genes, although they found a very low frequency of about 20 non-*Elysia* sequences (possibly contaminants) whereby they noted that "*None of these 20 ESTs, however, has a direct involvement in photosynthesis*". However, that did not stop them from concluding nonetheless that "*multiple lines of evidence that indicate that nuclear algal genes have been transferred*" (Pelletreau et al. 2011, pp. 1563, 1564). A problem with that conclusion is that the direct test of functionally relevant gene transfer (EST sequencing) comes up negative; all other data hinting at HGT in the slugs are indirect.

In the most recently published chapter of this story, Pierce et al. (2012) reported 98,000,000 reads and 8.8 billion base pairs of next-generation transcriptome sequencing data for *E. chlorotica*, a huge and valuable sequencing effort. They found many thousand reads that derive from the chloroplast genome, including 4,200 reads for the D1 protein of photosystem II alone. But among their 98 million reads, they only find 101 that match *Vaucheria* nuclear sequence data better than animal sequence data, corresponding to 52 transcripts. However, from that they conclude "*that a variety of functional algal genes have been transferred into the slug genome*" (Pierce et al. 2012, p. 3), even though evidence for expression of the collection of well over 500 nuclear-encoded genes that are required for functional plastids (Fig. 5) is altogether lacking. We disagree with their conclusion that their data represents evidence for HGT. In our view, the main finding of their study is that 98 million reads from *E. chlorotica* nucleic acid preparations produce 100 reads with undeniable similarity to *Vaucheria* nuclear genes; that finding suggests to us that their nucleic acid preparations used for sequencing were 99.9999 % free of contamination, because that is the frequency (one part per million) at which they find *Elysia* sequences. A purity of 99.9999 % is outstanding by any measure. But at the same time, the case for HGT to *Elysia* becomes a one in a million proposition.

Fig. 5 (continued) match with these first 50 genes or any *Acetabularia* gene. Three contigs of *P. ocellatus* (*black arrows*) had a match with one of the 500 most highly expressed *Arabidopsis* genes. One of these three (Rpl5, marked with *double asterisk*) is plastome encoded in chlorophytes and not nuclear encoded as in *Arabidopsis*. The other two are superoxide dismutase and a zinc finger protein. One match (*white arrowhead*) occurs in *E. timida*, it refers to ferritin. These three proteins/enzymes are ubiquitous

Future Research

If we accept the evidence from three laboratories showing that there is no horizontal gene transfer behind sacoglossan kleptoplasty (despite some interpretations of that evidence to the contrary), we have to ask again: what makes a slug photosynthetic? There are many unsolved problems concerning evolution, ecology, behaviour and genetics of these unique metazoan life forms. It starts with a founded hypothesis on the phylogenetic relationship of Sacoglossa and its sister taxon (Wägele et al. 2008; Händeler et al. 2009; Dinapoli and Klussmann-Kolb 2010; Jörger et al. 2011). This is a prerequisite to understanding when and why plastid incorporation has evolved, as well as the role of the food algae. At the moment, it seems that incorporation without digestion could have evolved once in the stemline of the Plakobranchoidea (Händeler et al. 2009), or alternatively multiple times independently. Available results, for example, on *C. ocellifera*, would support the latter hypothesis. This species, belonging to the Limapontioidea, shows a retention period of more than 60 days (Clark et al. 1981), in contrast to congeneric species [see Händeler et al. (2009)].

In the context of phylogeny, we can ask what the driving force behind the origin of long-term plastid retention was. Photosynthesis in slugs was certainly a by-product in evolution, fuelled by selective advantages (Ros and Marín 1990). Retaining or obtaining green coloration might have started as a means to becoming camouflaged (Clark et al. 1990; Rumpho et al. 2000; Wägele and Klussmann-Kolb 2005). Even shelled sacoglossans are green, although this colour is not derived from incorporated plastids. Additional input of energy by performing photosynthesis also provides an advantage for the food resources. These resources can be spared and food shortage (e.g. due to calcification) or even feeding on rare algal species becomes unproblematic (Marín and Ros 1989; Ros and Marín 1990; Teugels et al. 2008). Photosynthesis also helps the slug by supplementing the energetic demands of synthesizing defence compounds and possibly by facilitating nitrogen acquisition (Rumpho et al. 2006; Teugels et al. 2008). The ability to photosynthesize also helps relieve the energetic demands from reproductive behaviour, as Jensen (1987) demonstrated, when she observed reduced copulatory activity in starved *Ercolania nigra*, a non-retention form. It seems very likely that photosynthesis then also increases reproductive output.

We also have to ask whether the incorporation of plastids and subsequent ability to perform photosynthesis contributed to radiation within Sacoglossa. The number of sacoglossans is hardly known. Our own expeditions (H. Wägele) have revealed dozens of previously unknown sacoglossan species in need of formal description, and the same holds true for the collections of several colleagues. New molecular systematic analyses additionally revealed cryptic speciation (Carmona et al. 2011), and future analyses will show whether these new species are also characterized by their different and not yet investigated (photo-) biology. The few recently described species where photobiology was studied certainly raised the number of known photosynthetic slugs (Wägele et al. 2010b; Swennen 2011). However, there are

other factors to consider, for example chemical defence and protection against irradiance by uptake of secondary metabolites from food and/or their de novo synthesis (Gavagnin et al. 2000; Marín and Ros 2004).

Concerning the properties of the slug, we can ask whether they can regulate the uptake of chloroplasts that enable photosynthesis. Do they prefer certain plastids over others? Trench (1975) reported that *Elysia cauze* selected photosynthetic plastids for engulfment and rejected amylogenic ones. This would indicate a selective uptake of plastids from the same alga, depending on specific plastid features. Subsequently, we can ask about plastids from different algae and why they are not digested. Muscatine et al. (1975) and McLean (1976) have proposed that phagocytosis of plastids by symbiotic sacoglossans is followed by lysis of the enveloping host membrane, but the electron microscopic observations on enveloping membranes of sequestered plastids are still not completely clear, and even less so after plastid phagocytosis and incorporation for days up to months in STR and LTR slugs.

We also do not know the importance behind kleptoplast origin—a feature stressed by Evertsen and Johnsen (2009). Their studies showed that, in contrast to incorporated *Vaucheria* plastids in *E. chlorotica*, pigment proteins were not synthesized in *E. viridis* when they incorporate plastids from *C. fragile*. Therefore, the functionality of retained plastids could be investigated with regard to a slug's ability to synthesize pigments, proteins, lipids and starch. Recent success in molecular analyses of incorporated plastids by barcoding now provide the facility to clarify which food contributes to photosynthesis and which food does not (Händeler et al. 2010; Wägele et al. 2011).

Results taken from literature indicate that not only the plastid and its environment within the slug's cell counts but also the slug itself. Studies on ecology and adaptations in morphology and behaviour are also important. At least some species are able to reduce irradiance and therefore photodamage by phototactic behaviour and shading by parapodia (Schmitt and Wägele 2011). More studies on the influence of temperature and seasonality are needed, because their impact on photosynthetic performance seems to be very high [see Clark et al. (1981), Waugh and Clark (1986)].

We also know little concerning how much slug behaviour is influenced by the incorporation of plastids, or if their behaviour changes when enduring starvation. This seems very likely since reproductive efforts profit from photosynthesis (Jensen 1987; Middlebrooks et al. 2011); however, nothing is yet known about how this behaviour is triggered. Are there special photoreceptors as studies on eyeless *E. timida* specimens suggest (Rahat and Monselise 1979)?

Finally, what enables the slugs to perform photosynthesis for many months? Transcriptomic studies on three long-term retention forms—*Placobranchus ocellatus* (Wägele et al. 2011), *E. timida* (Wägele et al. 2011) and *E. chlorotica* (Pelletreau et al. 2011; Pierce et al. 2012) reject the hypothesis that HGT underlies plastid longevity [or should, see Pierce et al. (2012)]. Therefore, future analyses on plastid stability, protein stability and the slug's biochemical contribution should move to the fore in efforts to understand this beautiful, fascinating, and—among metazoans unique—symbiotic relationship.

Acknowledgements We wish to thank Katharina Händeler (Bonn), Gregor Christa (Bonn), Valerie Schmitt (Düsseldorf), Rainer Martin (Ulm) and Elise Lätz (Bonn) for discussions. The DFG (German Science Foundation) (Wa 618/8, Wa 618/12, SFBTr1) and the ERC (Networkorigins) provided financial support. Information on taxonomic status of algal species were taken from *AlgaeBase* (Guiry and Guiry 2012) and on sacoglossan species from the Seaslugforum (Rudman 2012) (http://www.seaslugforum.net/; searched on 12 February 2012). We follow the conventions of the International Code of Zoological/Botanical Nomenclature and include names of the authors who described the mentioned species and additionally cite the original work.

References

Agardh CA (1823) Species algarum rite cognitae, cum synonymis, differentiis specificis et descriptionibus succinctis. Volumen primum pars posterior. ex officina Berlingiana, Lundae, pp 399–531

Agardh CA (1873) Till algernes systematik. Nya bidrag. Lunds Universitets Års-Skrift. Afd Mathematik Naturvetenskap 9:1–71

Agosin M, Repetto Y (1963) Studies on the metabolism of *Echinococcus granulosus** – VII. Reactions of the tricarboxylic acid cycle in *E. granulosus* scolices. Comp Biochem Physiol 8:245–261

Alder J, Hancock A (1843) Notice on a British species of *Calliopaea* d'Orbigny, and on four new species of *Eolis* with observations on the development and structure of the nudibranchiate Mollusca. Ann Mag Nat Hist 12:233–238

Alder H, Hancock A (1862) Descriptions of a new genus and some new species of naked Mollusca. Ann Mag Nat Hist 10:261–265

Baba K (1955) Opisthobranchia of Sagami Bay, supplement. Iwanami Shoten, Tokyo, p 59

Brandley BK (1981) Ultrastructure of the envelope of *Codium australicum* (Silva) chloroplasts in the alga and after acquisition by *Elysia maoria* (Powell). New Phytol 89:679–686

Brandley BK (1984) Aspects of the ecology and physiology of *Elysia* cf. *furvacauda* (Mollusca: Sacoglossa). Bull Mar Sci 34:207–219

Burn RF (1958) Further Victorian Opisthobranchia. J Malacol Soc Aust 2:20–36

Carmona L, Malaquias MAE, Gosliner TM, Pola M, Cervera JL (2011) Amphi-Atlantic distributitions and cryptic species in sacoglossan sea slugs. J Molluscan Stud 77:401–412

Clark KB, Bussacca M (1978) Feeding specifity and chloroplast retention in four tropical Ascoglossa, with a discussion of the extent of chloroplast symbiosis and the evolution of the order. J Molluscan Stud 44:272–282

Clark KB, Jensen KR, Stirts HM, Fermin C (1981) Chloroplast symbiosis in a non-Elysiid mollusc, *Costasiella lilianae* Marcus (Hermaeidae: Ascoglossa (= Sacoglossa)): effects of temperature, light intensity, and starvation on carbon fixation rate. Biol Bull 160:42–54

Clark KB, Jensen KR, Stirts HM (1990) Survey for functional kleptoplasty among West Atlantic Ascoglossa (= Sacoglossa) (Mollusca: Opisthobranchia). Veliger 33:339–345

Curtis NE, Massey SE, Schwartz JA, Maugel TK, Pierce SK (2005) The intracellular, functional chloroplasts in adult sea slugs (*Elysia crispata*) come from several algal species, and are also different from those in juvenile slugs. Microsc Microanal 11:1194–1195

Curtis NE, Massey SE, Pierce SK (2006) The symbiotic chloroplasts in the sacoglossan *Elysia clarki* are from several algal species. Invert Biol 125:336–345

de Lamarck JBPA (1813) Sur les polypiers empâtés. Ann Mus Hist Nat, Paris 20:294–312, 370–386, 432–458

De Negri A, De Negri G (1876) Bericht A. und G. De Negri (Gass. chim.). Ber Deut Chem Gesell 9:84

Decaisne J (1842) Mémoire sur les corallines ou polypiers calcifères [la seconde partie du "Essais sur une classification des algues et des polypiers calcifères de Lamourous"]. Ann Sci Nat, Botanique, Sec Sér 18:96–128
Di Marzo V, Marín A, Vardaro RR, De Petrocellis L, Villani G, Cimino G (1993) Histological and biochemical bases of defense mechanisms in four species of Polybranchioidea ascoglossan molluscs. Mar Biol 117:367–380
Dinapoli A, Klussmann-Kolb A (2010) The long way to diversity – phylogeny and evolution of the Heterobranchia (Mollusca: Gastropoda). Mol Phylogenet Evol 55:60–76
Edmunds M (1963) *Berthelinia caribbea* n.sp., a bivalved gastropod from the Western Atlantic. Zool J Linn Soc 44:731–739
Er M (1961) Opisthobranch mollusks from California. Veliger 3(Suppl):1–85
Er M, Ev M (1967) American opisthobranch molluscs Part I, tropical American Opisthobranchs. Stud Trop Oceanogr Miami 6:1–256
Evertsen J (2006) Retention of photosynthetic chloroplasts in five sacoglossans (Mollusca: Opisthobranchia) from the Houtman Abrolhos Islands. Rec West Aust Mus Suppl 69:133–135
Evertsen J, Johnsen G (2009) In vivo and in vitro differences in chloroplast functionality in the two north Atlantic sacoglossans (Gastropoda, Opisthobranchia) *Placida dendritica* and *Elysia viridis*. Mar Biol 156:847–859
Evertsen J, Burghardt I, Johnsen G, Wägele H (2007) Retention of functional chloroplasts in some sacoglossans from the Indo-Pacific and Mediterranean. Mar Biol 151:2159–2166
Gavagnin M, Marín A, Mollo E, Crispino A, Villani G, Cimino G (1994) Secondary metabolites from Mediterranean Elysioidea: origin and biological role. Comp Biochem Physiol 108B:107–115
Gavagnin M, Mollo E, Montanaro D, Ortea J, Cimino G (2000) Chemical studies of Caribbean sacoglossans: dietary relationships with green algae and ecological implications. J Chem Ecol 26:1563–1578
Giles KL, Sarafis V (1972) Chloroplast survival and division in vitro. Nat New Biol 236:56–58
Giménez Casalduero F, Muniain C (2008) The role of kleptoplasts in the survival rates of *Elysia timida* (Risso 1818): (Sacoglossa: Opisthobranchia) during periods of food shortage. J Exp Mar Biol Ecol 357:181–187
Gould AA (1870) Report on the Invertebrata of Massachusetts. Second edition, comprising the Mollusca. Wright and Potter, Boston, p 524
Gould SB, Waller RR, McFadden GI (2008) Plastid evolution. Annu Rev Plant Biol 59:491–517
Grant BR, Borowitzka MA (1984a) Changes in structure of isolated chloroplasts of *Codium fragile* and *Caulerpa filiformis* in response to osmotic shock and detergent treatment. Protoplasma 120:155–164
Grant BR, Borowitzka MA (1984b) The chloroplasts of giant-celled and coenocytic algae: biochemistry and structure. Bot Rev 50:267–307
Graves DA, Gibson MA, Bleakney JS (1979) The digestive diverticula of *Alderia modesta* and *Elysia chlorotica*. Veliger 21:415–422
Green BJ, Li W-Y, Manhart JR, Fox TC, Summer EJ, Kennedy RA, Pierce SK, Rumpho ME (2000) Mollusc-algal chloroplast endosymbiosis. Photosynthesis, thylakoid protein maintenance, and chloroplast gene expression continue for many months in the absence of the algal nucleus. Plant Physiol 124:331–342
Green BJ, Fox TC, Rumpho ME (2005) Stability of isolated algal chloroplasts that participate in a unique mollusc/kleptoplast association. Symbiosis 40:31–40
Greene RW (1970) Symbiosis in sacoglossan opisthobranchs: functional capacity of symbiotic chloroplasts. Mar Biol 7:138–142
Guiry MD, Guiry GM (2012) *AlgaeBase*. World-wide electronic publication, National University of Ireland, Galway. http://www.algaebase.org; searched on 12 Feb 2012
Händeler K, Wägele H (2007) Preliminary study on molecular phylogeny of Sacoglossa and a compilation of their food organisms. Bonn Zool Beitr 3(4):231–254

Händeler K, Grzymbowski Y, Krug JP, Wägele H (2009) Functional chloroplasts in metazoan cells – a unique evolutionary strategy in animal life. Front Zool 6:28
Händeler K, Wägele H, Wahrmund U, Rüdinger M, Knoop V (2010) Slugs' last meals: molecular identification of sequestered chloroplasts from different algal origins in Sacoglossa (Opisthobranchia, Gastropoda). Mol Ecol Res 10:968–978
Hanten JJ, Pierce SK (2001) Synthesis of several light-harvesting complex I polypeptides is blocked by cycloheximide in symbiotic chloroplast in the sea slug, *Elysia chlorotica* (Gould): a case for horizontal gene transfer between alga and animal? Biol Bull 201:34–44
Hariot P (1889) Mission Scientifique du Cap Horn. 1882–188.3 Botanique I. Algues. Gauthier-Villars et fils, Paris, pp 3–109
van Hasselt JC (1824) In: Férussac A. Extrait d'une lettre du Dr. J.C. van Hasselt au Prof. van Swinderen sur mollusques de Java (traduit de l'Allgem. konst en letterbode, 1824, nos. 2,3,4) Tjuringe (ile Java) le 25 Mai 1823 (1). Bull Sci Nat Geol 3:237–248
Hawes CR (1979) Ultrastructural aspects of the symbiosis between algal chloroplasts and *Elysia viridis*. New Phytol 83:445–450
Hawes CR, Cobb AH (1980) The effects of starvation on the symbiotic chloroplasts in *Elysia viridis*: a fine structural study. New Phytol 84:375–379
Hinde R, Smith DC (1972) Persistence of functional chloroplasts in *Elysia viridis* (Opisthobranchia, Sacoglossa). Nat New Biol 239:30–31
Hinde R, Smith DC (1974) "Chloroplast symbiosis" and the extent to which it occurs in Sacoglossa (Gastropoda: Mollusca). Biol J Linn Soc 6:349–356
Hinde R, Smith DC (1975) The role of photosynthesis in the nutrition of the mollusc *Elysia viridis*. Biol J Linn Soc 7:161–171
Hirose E (2005) Digestive system of the Sacoglossan *Plakobranchus ocellatus* (Gastropoda: Opisthobranchia): light- and electron-microscopic observations with remarks on chloroplast retention. Zoolog Sci 22:905–916
Ireland C, Scheuer PJ (1979) Photosynthetic marine mollusks: in vivo ^{14}C incorporation into metabolites of the sacoglossan *Placobranchus ocellatus*. Science 205:922–923
Jensen KR (1980) A review of sacoglossan diets, with comparative notes on radular and buccal anatomy. Malac Rev 13:55–77
Jensen KR (1987) Effect of starvation on copulatory activity of *Ercolania nigra* (Lemche, 1935) (Mollusca, Opisthobranchia, Ascoglossa). Mar Behav Physiol 13:89–97
Jensen KR (1993a) Evolution of buccal apparatus and diet radiation in the Sacoglossa (Opisthobranchia). Boll Malacol 29:147–172
Jensen KR (1993b) Morphological adaptations and plasticity of radular teeth of the Sacoglossa (=Ascoglossa) (Mollusca: Opisthobranchia) in relation to their food plants. Biol J Linn Soc 48:135–155
Jensen KR (1994) Behavioural adaptations and diet specificity of sacoglossan opisthobranchs. Ethol Ecol Evol 6:87–101
Jesus B, Ventura P, Calado G (2010) Behaviour and a functional xanthophyll cycle enhance photoregulation mechanisms in the solar-powered sea slug *Elysia timida* (Risso, 1818). J Exp Mar Biol Ecol 395:98–105
Johnson MD (2011) Acquired phototrophy in ciliates: a review of cellular interactions and structural adaptations. J Eukaryot Microb 58:185–195
Jörger KM, Stöger I, Kano Y, Fukuda H, Knebelsberger T, Schrödl M (2011) On the origin of Acochlidia and other enigmatic euthyneuran gastropods, with implications for the systematics of Heterobranchia. BMC Evol Biol 10:323
Kawaguti S, Yamasu T (1965) Electron microscopy on the symbiosis between an elysioid gastropod and chloroplasts of a green alga. Biol J Okayama Univ 11:57–65
Klochkova TA, Han JW, Kim J-H, Kim KY, Kim GH (2010) Feeding specifity and photosynthetic activity of Korean sacoglossan mollusks. Algae 25:217–227

Koike K, Sekiguchi H, Kobiyama A, Takishita K, Kawachi M, Koike K, Ogata T (2005) A novel type of kleptoplastidy in *Dinophysis* (Dinophyceae): presence of haptophyte-type plastid in *Dinophysis mitra*. Protist 156:225–237
Lamouroux JVF (1816) Histoire des polypiers coralligènes flexibles, vulgairement nommés zoophytes. F. Poisson, Caen, pp 1–560
Lee JJ (2006) Algal symbiosis in larger foraminifera. Symbiosis 42:63–75
Lovén SL (1844) Om nordiska hafs-mollusker. Kongl Vetensk Acad Handl 1:48–53
Maeda T, Kajita T, Maruyama T, Hirano Y (2010) Molecular phylogeny of the Sacoglossa, with a discussion of gain and loss of kleptoplasty in the evolution of the group. Biol Bull 219:17–26
Maggs CA, Hommersand MH (1993) Seaweeds of the British Isles, vol 1, Rhodophyta. Part 3A. Ceramiales. HMSO, London, pp 1–444
Marín A, Ros J (1989) The chloroplast-animal association in four Iberian sacoglossan opisthobranchs: *Elysia timida*, *Elysia translucens*, *Thuridilla hopei*, and *Bosellia mimetica*. Sci Mar 53:429–440
Marín A, Ros J (1992) Dynamics of a peculiar plant-herbivore relationship: the photosynthetic ascoglossan *Elysia timida* and the chlorophycean *Acetabularia acetabulum*. Mar Biol 112:677–682
Marín A, Ros J (1993) Ultrastructural and ecological aspects of the development of chloroplast retention in the sacoglossan gastropod *Elysia timida*. J Molluscan Stud 59:95–104
Marín A, Ros J (2004) Chemical defenses in sacoglossan opisthobranchs: taxonomic trends and evolutive implications. Sci Mar 68:227–241
Martin W, Cerff R (1986) Prokaryotic features of a nucleus encoded enzyme: cDNA sequences for chloroplast and cytosolyic glyceraldehyde-3-phosphate dehydrogenases from mustard (*Sinapis alba*). Eur J Biochem 159:323–331
Martin W, Herrmann RG (1998) Gene transfer from organelles to the nucleus: how much, what happens, and why? Plant Physiol 118:9–17
Martin W, Schnarrenberger C (1997) The evolution of the Calvin cycle from prokaryotic to eukaryotic chromosomes: a case study of functional redundancy in ancient pathways through endosymbiosis. Curr Genet 32:1–18
Martin W, Brinkmann H, Savona C, Cerff R (1993) Evidence for a chimaeric nature of nuclear genomes: eubacterial origin of eukaryotic glyceraldehyde-3-phosphate dehydrogenase genes. Proc Natl Acad Sci USA 90:8692–8696
Martin W, Stoebe B, Goremykin V, Hansmann S, Hasegawa M, Kowallik KV (1998) Gene transfer to the nucleus and the evolution of chloroplasts. Nature 393:162–165
Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, Leister D, Stoebe B, Hasegawa M, Penny D (2002) Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proc Natl Acad Sci USA 99:12246–12251
McLean N (1976) Phagocytosis of chloroplasts in *Placida dendritica* (Gastropoda: Sacoglossa). J Exp Zool 197:321–330
Mereschkowsky C (1905) Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. Biol Centralbl 25:593–604 [English translation in Martin W, Kowallik KV (1999) Eur J Phycol 34: 287–295]
Middlebrooks ML, Pierce SK, Bell SS (2011) Foraging behavior under starvation conditions is altered via photosynthesis by the marine gastropod. *Elysia clarki*. PLoS One 6
Monselise EBI, Rahat M (1980) Photobiology of *Elysia timida* (Mollusca: Opisthobranchia): observations in the sea. Isr J Zool 29:125–128
Montagu G (1804) Description of several marine animals found on the south coast of Devonshire. Trans Linn Soc 11:1–26
Montagu G (1815) An account of some new and rare British shells and animals. Trans Linn Soc Lond 11:179–204
Mørch OAL (1863) Contributions a la faune malacologique des Antilles Danoises. J Conchol 11:21–43

Mujer CV, Andrews DL, Manhart JR, Pierce SK, Rumpho ME (1996) Chloroplast genes are expressed during intracellular symbiotic association of *Vaucheria litorea* plastids with the sea slug *Elysia chlorotica*. Proc Natl Acad Sci USA 93:12333–12338

Müller OF (1774) Vermium terredtrium et fluviatilium, vol 2. Havniae & Lipsiae

Muscatine L, Pool RR, Trench RK (1975) Symbiosis of algae and invertebrates: aspects of the symbiont surface and the host-symbiont interface. Trans Am Microsc Soc 94:450–469

Pelletreau KN, Bhattacharya D, Price DC, Worful JM, Moustafa A, Rumpho ME (2011) Sea slug kleptoplasty and plastid maintenance in a metazoan. Plant Physiol 155:1561–1565

Pierce SK, Biron RW, Rumpho ME (1996) Endosymbiotic chloroplasts in molluscan cells contain proteins synthesized after plastid capture. J Exp Biol 199:2323–2330

Pierce SK, Massey SE, Hanten JJ, Curtis NE (2003) Horizontal transfer of functional nuclear genes between multicellular organisms. Biol Bull 204:237–240

Pierce SK, Curtis NE, Massey SE, Bass AL, Karl SA, Finney CM (2006) A morphological and molecular comparison between *Elysia crispata* and a new species of kleptoplastic sacoglossan sea slug (Gastropoda: Opisthobranchia) from the Florida Keys, USA. Moll Res 26:23–38

Pierce SK, Curtis NE, Hanten JJ, Boerner SL, Schwartz JA (2007) Transfer, integration and expression of functional nuclear genes between multicellular species. Symbiosis 43:57–64

Pierce SK, Curtis NE, Schwartz JA (2009) Chlorophyll *a* synthesis by an animal using transferred algal nuclear genes. Symbiosis 49:121–131

Pierce SK, Fang X, Schwartz JA, Jian X, Zhao W, Curtis NE, Kocot K, Yang B, Wang J (2012) Transcriptomic evidence for the expression of horizontally transferred algal nuclear genes in the photosynthetic sea slug *Elysia chlorotica*. Mol Biol Evol 29:1545–1556

Pruvot-Fol A (1957) Diagnose dune Elysie peut-être nouvelle de la Méditerranée. Bull Mus Nat d'Hist Natur 29:337–339

Rahat M, Monselise EBI (1979) Photobiology of the chloroplast hosting mollusc *Elysia timida* (Opisthobranchia). J Exp Biol 79:225–233

Risso A (1818) Mémoire sur quelques gastéropodes nouveaux, nudibranches et tectibranches observés dans la Mer de Nice. J Phys Chim Hist Nat Arts 87:368–377

Ros J, Marín A (1990) Adaptive advantages of the "symbiosis" between algal chloroplasts and sacoglossan molluscs. Oecol Aquat 10:271–298

Rudman WB (2012) Seaslugforum. http://www.seaslugforum.net/; World-wide electronic publication, Australian Museum; searched on 12 Feb 2012

Rumpho ME, Summer EJ, Manhart JR (2000) Solar-powered sea slugs. Mollusc/algal chloroplast symbiosis. Plant Physiol 123:29–38

Rumpho ME, Summer EJ, Green BJ, Fox TC, Manhart JR (2001) Mollusc/algal chloroplast symbiosis: how can isolated chloroplasts continue to function for months in the cytosol of a sea slug in the absence of an algal nucleus? Zoology 104:303–312

Rumpho ME, Dastoor FP, Manhart JR, Lee J (2006) The kleptoplast. In: Wise RR, Hoober JK (eds) The structure and function of plastids. Springer, The Netherlands, pp 451–473

Rumpho ME, Worful JM, Lee J, Kannan K, Tylor MS, Bhattacharya D, Moustafa A, Manhart JR (2008) Horizontal gene transfer of the algal nuclear gene psbO to the photosynthetic sea slug *Elysia chlorotica*. Proc Natl Acad Sci USA 105:17867–17871

Rumpho ME, Pochareddy S, Worful JM, Summer EJ, Bhattacharya D, Pelletreau KN, Tylor MS, Lee J, Manhart JR, Soule KM (2009) Molecular characterization of the Calvin Cycle enzyme phosphoribulokinase in the stramenopile alga *Vaucheria litorea* and the plastid hosting mollusc *Elysia chlorotica*. Mol Plant 2:1384–1396

Rumpho ME, Pelletreau KN, Moustafa A, Bhattacharya D (2011) The making of a photosynthetic animal. J Exp Biol 214:303–311

Schimper AFW (1883) Über die Entwicklung der Chlorophyllkörner und Farbkörper. Bot Zeitung 41:105–114

Schmitt V, Wägele H (2011) Behavioral adaptations in relation to long-term retention of endosymbiotic chloroplasts in the sea slug *Elysia timida* (Opisthobranchia, Sacoglossa). Thalassas 27:225–238

Schöttler U, Bennet EM (1991) Annelids. In: Bryant C (ed) Metazoan life without oxygen. Chapman and Hall, London, UK, pp 165–185
Schwartz JA, Curtis NE, Pierce SK (2010) Using algal transcriptome sequences to identify transferred genes in the sea slug, *Elysia chlorotica*. Evol Biol 37:29–37
Silva PC (1952) A review of nomenclatural conservation in the algae from the point of view of the type method. Univ Calif Publ Bot 25:241–323
Simroth H (1895) Neuere Arbeiten über Opisthobranchien. Zool Centralbl 2:513–515
Stoecker DK, Johnson MD, de Vargas C, Not F (2009) Acquired phototrophy in aquatic protists. Aquat Microb Ecol 57:279–310
Swennen C (2011) Large mangrove-dwelling *Elysia* species in Asia, with descriptions of two new species (Gastropoda: Opisthobranchia: Sacoglossa). Raffl Bull Zool 59:29–37
Taylor DL (1968) Chloroplasts as symbiotic organelles in the digestive gland of *Elysia viridis* (Gastropoda: Opisthobranchia). J Mar Biol Assoc UK 48:1–15
Taylor DL (1971) Symbiosis between the chloroplasts of *Griffithsia flosculosa* (Rhodophyta) and *Hermaea bifida* (Gastropoda: Opisthobranchia). Pubbl Staz Zool Nap 39:116–120
Teugels B, Bouillon S, Veuger B, Middelburg JJ, Koedam N (2008) Kleptoplasts mediate nitrogen acquisition in the sea slug *Elysia viridis*. Aquat Biol 4:15–21
Thompson TE, Jarman GM (1989) Nutrition of *Tridachia crispata* (Mörch) (Sacoglossa). J Molluscan Stud 55:239–244
Timmis JN, Ayliffe MA, Huang CY, Martin W (2004) Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nat Rev Genet 5:123–135
Trench RK (1969) Chloroplasts as functional endysmbionts in the Mollusc *Tridachia crispata* (Bergh), (Opisthobranchia, Sacoglossa). Nature 222:1071–1072
Trench RK (1973) Further studies on the mucoploysaccharide secreted by the pedal gland of the marine slug *Tridachia crispata* (Opisthobranchia, Sacoglossa). Bull Mar Sci 23:299–312
Trench RK (1975) Of "leaves that crawl", functional chloroplasts in animal cells. Symp Soc Exp Biol 29:229–265
Trench RK, Smith DC (1970) Synthesis of pigment in symbiotic chloroplasts. Nature 227:196–197
Trench RK, Greene RW, Bystrom BG (1969) Chloroplasts as functional organelles in animal tissues. J Cell Biol 42:404–417
Trench ME, Trench RK, Muscatine L (1970) Utilization of photosynthetic products of symbiotic chloroplasts in mucus synthesis by *Placobranchus ianthobapsus* (Gould), Opisthobranchia, Sacoglossa. Comp Biochem Physiol 37:113–117
Trench RK, Boyle JE, Smith DC (1973a) The association between chloroplasts of *Codium fragile* and the mollusc *Elysia viridis* I. Characteristics of isolated *Codium* chloroplasts. Proc R Soc Lond B Biol Sci 184:51–61
Trench RK, Boyle JE, Smith DC (1973b) The association between chloroplasts of *Codium fragile* and the mollusc *Elysia viridis* II. Chloroplast ultrastructure and photosynthetic carbon fixation in *E. viridis*. Proc R Soc Lond B Biol Sci 184:63–81
Trench RK, Boyle JE, Smith DC (1974) The association between chloroplasts of *Codium fragile* and the mollusc *Elysia viridis* III. Movement of photosynthetically fixed 14C in tissues of intact living *E. viridis* and in *Tridachia crispata*. Proc R Soc Lond B Biol Sci 185:453–464
Verany M (1853) Catalogue des Mollusques cephalopodes, pteropodes, Gasteropodes nudibranches, etc.... des environs de Nice. J Conchiol 4:375–392
Vieira S, Calado G, Coelho H, Serôdio J (2009) Effects of light exposure on the retention of kleptoplastic photosynthetic activity in the sacoglossan mollusc *Elysia viridis*. Mar Biol 156:1007–1020
Wägele H, Johnsen G (2001) Observations on the histology and photosynthetic performance of "solar-powered" opisthobranchs (Mollusca, Gastropoda; Opisthobranchia) containing symbiotic chloroplasts or zooxanthellae. Org Divers Evol 1:193–210
Wägele H, Klussmann-Kolb A (2005) Opisthobranchia (Mollusca, Gastropoda) – more than just slimy slugs. Shell reduction and its implications on defence and foraging. Front Zool 2:1–18

Wägele H, Klussmann-Kolb A, Vonnemann V, Medina M (2008) Heterobranchia I, the Opisthobranchia. In: Ponder W, Lindbergh W (eds) Phylogeny and evolution of the Mollusca. University of California Press, Berkeley, CA, pp 383–406

Wägele H, Raupach MJ, Burghardt I, Grzymbowski Y, Händeler K (2010a) Solar powered seaslugs (Opisthobranchia, Gastropoda, Mollusca): incorporation of photosynthetic units: a key character enhancing radiation? In: Glaubrecht M (ed) Evolution in action. Springer, Berlin, pp 263–283

Wägele H, Stemmer K, Burghardt I, Händeler K (2010b) Two new sacoglossan sea slug species (Opisthobranchia, Gastropoda): *Ercolania annelyleorum* sp. nov. (Limapontioidea) and *Elysia asbecki* sp. nov. (Plakobranchoidea) with notes on anatomy, histology and biology. Zootaxa 2676:1–28

Wägele H, Deusch O, Händeler K, Martin R, Schmitt V, Christa G, Pinzger B, Gould SB, Dagan T, Klussmann-Kolb A, Martin W (2011) Transcriptomic evidence that longevity of acquired plastids in the photosynthetic slugs *Elysia timida* and *Plakobranchus ocellatus* does not entail lateral transfer of algal nuclear genes. Mol Biol Evol 28:699–706

Waugh GR, Clark KB (1986) Seasonal and geographic variation in chlorophyll level of *Elysia tuca* (Ascoglossa: Opisthobranchia). Mar Biol 92:483–487

Weaver S, Clark KB (1981) Light intensity and color preferences of five ascoglossan (=sacoglossan) molluscs (Gastropoda: Opisthobranchia): a comparison of chloroplast-symbiotic and aposymbiotic species. Mar Behav Physiol 7:297–306

Yamamoto YY, Yusa Y, Yamamoto S, Hirano Y, Hirano Y, Motomura T, Tanemura T, Obokata J (2009) Identification of photosynthetic sacoglossans from Japan. Endocyt Cell Res 19:112–119

Index

A
Amitochondriate eukaryotes, 9–10
Arabidopsis thaliana, 60–61

B
β-Barrel Assembly Machinery (BAM), 87
Bigelowiella natans, 207–209

C
Calvin–Benson cycle, 59, 60
Carbon-concentrating mechanism (CCM), 137
CCAC 0185 plastid genome, 160, 161
Chlamydiae, 71–72
Chlamydomonas reinhardtii, 43, 92, 95, 96, 98
Chromalveolate hypothesis, 57, 190–191
Chromatophores, 154
Chromera velia
  alveolate puzzle, 194
  apicomplexan lineage, 193
  chromalveolate hypothesis, 190–191
  chromerids, 194
  discovery of, 187–189
  morphology and life cycle stages
    Alveolata super group, 185
    alveolin localization, 187, 188
    apicomplexans, 185, 186
    asexual cell division, 187
    micropore, 185
    short-lived flagellated stage, 186
    symbiotic dinoflagellates, 185
    Toxoplasma gondii, 187
    ultrastructural analysis, 186
  phylogenomic position, 189
  phylogenomic relationship of, 194
  plastid bio-and photochemistry, 191–193
Core dinoflagellates
  cell organelles evolution, 241
  diatom-derived
    establishment of pennate, 262–263
    eyespots, 268
    general characteristics of, 259–260
    genomics of, 264–266
    independent replacements, 263–264
    monophyletic *vs.* polyphyletic origins, 262
    putative protein import, 266–268
    ultrastructure and origin of, 260–262
  dinophysis (*see* Dinophysis)
  feeding mechanisms in
    myzocytosis, 239–240
    pallium feeding, 240
    phagocytosis, 239
  general characteristics, 235
  Kareniaceae dinoflagellates
    fucoxanthin plastids (*see* Fucoxanthin dinoflagellates)
    general characteristics of, 269
    genome and transformation, 271–272
    haptophyte origin evidence for, 269–270
  kleptoplastids, 247–248
  mixotrophy, 239
  peridinin-containing plastid, 243, 245–246
  photosynthetic entities, 241, 243–244
  phylogeny of, 235, 236, 238
  ultrastructure of, 235, 237

W. Löffelhardt (ed.), *Endosymbiosis*, DOI 10.1007/978-3-7091-1303-5,
© Springer-Verlag Wien 2014

Core dinoflagellates (*cont.*)
unusual dinoflagellate plastids, 246–247
Cryptomonad and chlorarachniophyte algae. *See* Nucleomorph genomes
Cryptophyte-derived chromatophores, dinophysis
evidence for, 250
genomic evidence, 254
indirect Myrionecta rubra-mediated route, 250–251
molecular data, kleptoplastidy, 253–254
ultrastructural data support, 252–253
Cyanidioschyzon merolae, 57
Cyanobacterial ancestor, 47
Cyanophora paradoxa, 43, 61
CCM, 137
genome data, 137
muroplasts, 144–146
peptidoglycan biosynthesis
compartmentalization process, 141
lysozyme family protein, 143
murG, 143
N-acetylputrescine, 141
periplasmic proteins, 141
periplasmic space, 141
plantae lineages, 136
plantae monophyly, 138–141
primary plastids, 136

D
Diatom-derived dinoflagellates
establishment of pennate, 262–263
eyespots, 268
general characteristics of, 259–260
genomics of, 264–266
independent replacements, 263–264
monophyletic *vs*. polyphyletic origins, 262
putative protein import, 266–268
ultrastructure and origin of, 260–262
Dicarboxylate translocators (DiTs), 67
Dinophysis
cryptophyte-derived chromatophores
evidence for, 250
genomic evidence, 254
indirect Myrionecta rubra-mediated route, 250–251
molecular data, kleptoplastidy, 253–254
ultrastructural data support, 252–253
general characteristic of, 248–249
kleptoplastids
longevity, 258
nuclear-encoded plastid proteins, 254–257
outer envelope membrane, 257–258
plastid replacement, 258–259
DiTs. *See* Dicarboxylate translocators (DiTs)

E
Endosymbiotic gene transfer (EGT), 41–43, 46, 58, 153, 160–161, 208
Eukaryotic origin
genome sequences, 4
hydrogen hypothesis (*see* Hydrogen hypothesis)
phylogenetic methodology, 4
symbiogenesis, 8, 10
symbiotic relationships, 4–5

F
Floridean starch, 61
Fucoxanthin dinoflagellates
bipartite N-terminal targeting signals, 274
mode of acquisition, 273–274
plastid replacement
FVAPmotif, 277
GAPDH, plastid-targeted isoform, 276
nonphotosynthetic functions, 277
oxygen-evolving enhancer 1 (PsbO) protein, 274
protein targeting mechanisms, 273–274

G
Galdieria sulphuraria, 60, 64
Glc6-P/phosphate translocator (GPT), 64
Glutamine synthetase/glutamate synthase cycle (GS/GOGAT), 67
Glycogen debranching enzymes, 115
Glycogen–starch metabolism
ADP-glucose, 115, 116
amylopectin, 114
amylose, 114
archaeplastidal starch metabolism network, 114
cyanobacterial and host partners, 113
eukaryotic glycogen synthases and phosphorylases, 116
glycogen, 114, 115
glycogen debranching enzymes, 115
MOS metabolism, 115
purine/pyrimidine nucleotide sugars, 115
starch granule, 114
storage α-glucans, 114
GPT. *See* Glc6-P/phosphate translocator (GPT)
Green-sulfur bacteria, 54
Guillardia theta, 207–209

H
Hemiselmis andersenii
Holospora, 5
Horizontal gene transfer (HGT), 41, 56
Hydrogen hypothesis
  amitochondriate eukaryotes, 10
  archaeal methanogens, 10
  archaea-related eukaryotic genes, 11–12
  cytoplasmic ribosomes, 11
  homocentric theories, 13
  LBA, 12
  phagocytosis and endosymbiosis, 12–13
  phylogenetic trace, α-proteobacterial, 10
  RNA polymerases, 11
  well-conserved eukaryotic genes, 12

J
Jam-packed genomes, 201–204

K
Kareniaceae dinoflagellates
  fucoxanthin (*see* Fucoxanthin dinoflagellates)
  general characteristics of, 269
  genome and transformation, 271–272
  haptophyte origin evidence for, 269–270
Kleptoplastids, dinophysis
  longevity, 258
  nuclear-encoded plastid proteins, 254–257
  outer envelope membrane, 257–258
  plastid replacement, 258–259

L
Lateral gene transfer (LGT), 41
Long-branch-attraction (LBA), 12

M
Malto-oligosaccharides (MOS), 115
Metabolic symbiosis, 8, 9
Metabolite transporters
  carbon metabolism
    autotrophic plastids, 59
    Calvin–Benson cycle, 59, 60
    Chloroplastida, 66
    glaucophytes, 65–66
    glucose transporter, 66
    maltose, 66
    MEX protein, 66
    3-PGA, 59–60
    phosphate translocators, 62–64
    starch metabolism, 61–62
    triose phosphates, 60–61
    Viridiplantae, 60
  nitrogen metabolism
    ammonia, 66–67
    ATP, cyanobiont, 69–70
    cytosolic nitrate reductase, 67
    DiTs, 67
    GS/GOGAT system, 67
    nitrogen fixation, 68
    plastidic nitrite reductase, 67
  plastidic translocators, 70–72
  plastid stroma and cytosol, 59
Mitochondrial
  origin
    archezoan (SET) hypothesis, 6–8
    genome sequences, 4
    hydrogen hypothesis (*see* Hydrogen hypothesis)
    phylogenetic methodology, 4
    α-proteobacteria, 5, 7–8
    Rickettsia-like bacteria, 6
  protein import pathways
    ancestral endosymbiont, 20
    Arabidopsis thaliana, 23
    bacterial membrane protein, 21
    BAM complex, 22
    diverse model organisms, 20
    evolution of, 20
    host-symbiont relationship, 20
    mtHsp70, 24
    organelle-encoded gene products, 21
    organelle evolution (*see* Organelle evolution, mitochondrial)
    Oxa1 protein, 22
    OXA translocase, 21
    periplasmic protein, 21
    SAM complex, 22, 26
    SecYEG translocases function, 21, 22
    TIM22, 23
    TIM23 complex, 23, 24
    TOM complex, 25
    yidC2, 22
    YidC translocases function, 21
Myzocytosis, 239–240

N
Nostoc/Azolla symbiosis, 42
Nucleomorph genomes
  Bigelowiella natans, 207–209
  convergent genome evolution, 207
  cryptomonad nucleomorph, 200
  features of, 198, 199

Nucleomorph genomes (*cont.*)
Guillardia theta, 207–209
jam-packed genomes, 201–204
molecular biology use, 200
ORFan genes, 205–206
plastid-associated genes, 204–205
reductive evolution, 208
secondary endosymbionts, 199
"The Rule of Three,"201
Nucleoside triphosphate transporter (NTT), 69–70
Nucleotide sugar transporter (NST), 62–63

O
Omp85 proteins, 87
ORFan genes, 205–206
Organelle evolution, mitochondrial
bacterial SecYEG secretion, 29
cryptic signals, 30
import proteins and machines, 26, 27
protein sorting machinery, 26, 28
Tat pathway, 29
Tom70 function, 27–28
trypanosomatids and microsporidia, 28
Oxygenic photosynthesis, 55

P
Pallium feeding, 240
Paulinella chromatophora, 46, 72, 122, 151–162
Peptidoglycan, 141
Phagocytosis, 239
Phosphoenolpyruvate/phosphate translocator (PPT), 64
3-phosphoglyceric acid (3-PGA), 59–60
Photosynthetic *Paulinella, Synechococcus*-like cyanobacterium
alpha-cyanobacterial origin, 155
blue–green chromatophores, 154
concatenated plastid proteins, 155
EGT, 160–162
heterotrophic sister taxa, 156–157
plastid genomes, 159–160
postendosymbiotic speciation, 157–159
primary plastids origin, 152–153
WH5701, 156
Phycobilisomes, 56–57
Phylogenetic tree, 11
Plantae plastid, 152–153
Plastid
cyanophora paradoxa, 93
cytoplasmically produced proteins, 88
inner envelope translocon, 99–100
N-terminal signals, 89
outer envelope translocon
Chlamydomonas reinhardtii, 95
Cyanophora paradoxa, 98–99
DnaJ-type proteins, 98
general translocation pathway, 97
Omp85 homolog Toc75-III, 94
phylogenetic analysis, 98
POTRA domains, 94, 96
TOC, 94
Toc64, 97
Toc159 and Toc34, 96
Toc75 protein family, 94, 95
protein import (*see* Protein import)
protein insertion machinery, 100–101
protoplastid, 88–89
storage polysaccharide metabolism
ADP-glucose, 127–130
cyanobiont, 113
glucan polymerization, 129–131
glycogen–starch metabolism (*see* Glycogen–starch metabolism)
oxygenic photosynthesis, 112
phototrophy and diazotrophy incompatibility, 124, 126–127
starch metabolism (*see* Starch metabolism, Archaeplastida)
stromal proteins, 89
Toc75, 90–92
transit peptide, 89, 93
vesicle transport system, 88–89
POlypeptide–TRansport-Associated (POTRA) domains, 94
Preprotein transport
apicomplexa, 216
chlorarachniophytes, 216
cryptophytes, 216
ER membrane, 218–219
haptophytes, 216
heterokontophytes, 216
N-terminal targeting signals, 222–223
nucleus-encoded membrane proteins, 223
outermost membrane crossing, 220–221
plastid membrane, 219
PPM, 219
proteins integration, outermost membrane, 221–222
red algal origin, 216, 217
Primary endosymbiosis
anoxygenic photosynthetic bacteria, 54
ATP synthesis and reduction, 55
cyanobacterial genomes code, 57

cyanobacterium, 56
glaucophytes, 56
green algae and land plants (chlorophyta), 57
oxygenic photosynthesis, 55
photoautotrophic eukaryotes, 56
photoautotrophic organisms, 54
plastid-containing eukaryotes, 58
plastid-containing modern algae, 56
plastidic metabolic pathways, 58
red algae, 56–57
Protein import
PPC *vs.* stroma targeting, 223–225
preprotein transport
apicomplexa, 216
chlorarachniophytes, 216
cryptophytes, 216
ER membrane, 218–219
haptophytes, 216
heterokontophytes, 216
N-terminal targeting signals, 222–223
nucleus-encoded membrane proteins, 223
outermost membrane crossing, 220–221
plastid membrane, 219
PPM, 219
proteins integration, outermost membrane, 221–222
red algal origin, 216, 217
recycled and modified, 225–226
Protein translocation, chloroplasts envelope
general translocation pathway, 82
land plants, 82–85
plastids (*see* Plastid)
prokaryotic membranes
BAM, 87
cyanobacterial ancestor, 87
gram-negative bacteria, 86
membrane-embedded transport complexes, 85
protein targeting, 86
proteobacterial species, 85
SEC and TAT, 85–86
trigger factor, 86–87
TOC complex, 82
transit peptides, 82
Purple photosynthetic bacteria, 54

R

Red algae (rhodophytes), 56–57
Rhopalodia gibba, spheroid body
chromatophores, 176
cytoplasm, 168, 169
electron microscopy, 168, 169
genomes
Cyanothece sp. ATCC 51142, 172
DNA fosmid library, 172
DNA repair, 174–175
G/C content, 174
genome size, 172–173
mobile elements, 174
pseudogenes, 173
shotgun sequencing, 172
nitrogenase activity, 169–170
nitrogen fixing activity, 170
origin, 170–171
pyrenoids, 168
Rhopalodiaceae family, 168
secondary endosymbiont, 168
UCYN-A group, 175
RuBP carboxylase/oxygenase (Rubisco), 45, 46, 59, 67, 144, 145, 155, 190, 225, 272, 276, 310, 314

S

Secondary endosymbiosis, 57
Serial endosymbiosis hypothesis, 40
Serial endosymbiosis theory (SET), 7–8
Single primary endosymbiotic event
archaebacterium, 40
Archaeplastida, 40
cyanobacterial ancestor, 47
EGT, 41–43
endosymbiotic aerobic bacteria, 40
phylogenetic analysis, 44–45
plant lineage, 40, 41
plastid genomes, 43, 44
plastid protein import apparatus, 45–47
α-proteobacterium, 40
serial endosymbiosis hypothesis, 40
"shopping bag" model, 48
Sorting and Assembly Machinery (SAM), 22
Starch metabolism, Archaeplastida
ADP-glucose, 121, 123
Chloroplastida, 122
early glycogen and late starch accumulation stages, 124, 125
eukaryotic glycogen metabolism, 120
GBSSI, 118, 121
glaucophyta, 118
Glaucophyta and Rhodophyceae, 122
glycosyl nucleotide, 118
GWD/PWD enzymes, 119–120, 123

Starch metabolism (*cont.*)
parasites and symbionts, 122
phylogenetic origin, 116–118, 121
plastidial storage polysaccharide, 116
red algae and glaucophytes, 116
Rhodophyceae, 119
starch synthases, 120
UDP-glucose, 121
Stroma-targeting peptide (STP), 42–43

T
Toxoplasma gondii, 187
Triose phosphate/phosphate translocator (TPT), 60–61
Twin-Arginine-dependent Translocon (TAT), 85–86

U
UhpC homologues, 65

X
Xylulose-5-phosphate/phosphate translocator (XPT), 64